FINAL REPORT

Spatial Statistical Model and Optimal Survey Design for Rapid Geophysical Characterization of UXO Sites

SERDP Project CU-1201

G. Ostrouchov, W. E. Doll, D. A. Wolf, and L. P. Beard
Oak Ridge National Laboratory
Oak Ridge, Tennessee

M. D. Morris
Iowa State University
Ames, Iowa

D. K. Butler
U.S. Army Engineer Research and Development Center (ERDC)
Vicksburg, Mississippi

July, 2003

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **JUL 2003**	2. REPORT TYPE	3. DATES COVERED **00-00-2003 to 00-00-2003**
4. TITLE AND SUBTITLE **Spatial Statistical Model and Optimal Survey Design for Rapid Geophysical Characterization of UXO Sites**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Oak Ridge National Laboratory,Oak Ridge,TN,37831**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT **Same as Report (SAR)**	18. NUMBER OF PAGES **99**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

EXECUTIVE SUMMARY

Unexploded ordnance (UXO) surveys encompass large areas, and the cost of surveying these areas can be high. Enactment of earlier protocols for sampling UXO sites (SiteStats/GridStats) have shown the shortcomings of these procedures and led to a call for development of scientifically defensible statistical procedures for survey design and analysis. This project is one of three funded by SERDP to address this need.

The problem is a very complicated one statistically, with a need to develop new approaches in survey design and data analysis. It became evident to our team at an early stage in this project that all of the statistical tools that are needed to address this problem were not 'on the table,' and that some of the key questions had not been previously solved mathematically. It is critical that a new protocol for design and analysis of UXO surveys be able to address: (1) the idiosyncrasies of different sites in terms of their geology, ordnance types, topography, vegetation, and extent of background knowledge; (2) the uncertainties in performance of different types of instrumentation and instrument platforms that are available, or are becoming available for UXO surveys; (3) the distinctions between UXO contamination, which occurs at discrete points, and chemical contamination, which has a more continuous distribution; (4) the opportunities to interrogate the site through sequential surveys; and (5) the regulatory and public-involvement environment in which these surveys are typically performed. The protocol must allow for changes in these factors that may result from technological advances. A protocol that neglects some or all of these issues may never be suitable for routine use.

An alternative approach is to develop statistical tools that are appropriate under artificially simplistic settings, and to expand these tools over time to accommodate situations that are more and more realistic. We have chosen not to pursue this approach because we felt that the 'top down' approach was less likely to encounter insurmountable obstacles.

Statistically based methodologies are being used, and should be used, to efficiently determine the extent of UXO contamination by optimizing locations and geographic extent for surveys, defining how to conduct excavation, and developing procedures to incorporate survey data into decision making. These methods may also be used to prioritize areas, compare different clearance approaches, and to estimate costs for different land uses (i.e., for different specified levels of clearance).

Statistical methods are used to reduce the extent of data acquisition required to characterize large areas. Our protocol has led to development of statistically valid tools that can be used to support management decisions by providing maps of estimated contamination and associated probabilities and uncertainties at different stages of the characterization process. Such tools enable decisions that lead to more appropriate and cost effective remediation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AGL	Above Ground Level
ASR	Archive Search Report
AUC	Area Under the ROC Curve
BBR	Badlands Bombing Range
CWM	Chemical Warfare Material
DOAM	Distributions On A Map
EM	Electromagnetic
ESTCP	Environmental Security Technology Certification Program
FPR	False Positive Rate
FUDS	Formerly Used Defense Site
GIS	Geographical Information System
GUI	Graphical User Interface
INPR	Inventory Project Reports
LDA	Linear Discriminant Analysis
LG	Log-Gaussian
MAGMOD	Magnetic Modeling code (Butler et al., 1998)
MCMC	Markov Chain Monte Carlo
MTADS	Multisensor Towed Array Detection System
OE	Ordnance and Explosive
OIM	Ordnance Intensity Map
ORAGS	Oak Ridge Airborne Geophysical System
ORNL	Oak Ridge National Laboratory
pAUC	Partial Area Under the ROC Curve
ROC	Receiver Operator Characteristic
SERDP	Strategic Environmental Research and Development Program
SS/GS	SiteStats / GridStats
TDEM	Time-Domain Electromagnetic Induction
TFM	Total Field Magnetic
TIM	Target Intensity Map
TPR	True Positive Rate
USAESCH	U. S. Army Engineering and Support Center - Huntsville
UXO	Unexploded ordnance

GLOSSARY

Classification: Identification of the category or group to which an individual or object belongs on the basis of observed characteristics.

Euclidean Distance: A distance metric. The Euclidean distance between two points is defined as the square root of the sums of squares of coordinate differences.

Linear Discriminant Analysis: A statistical technique for classification by means of multivariate quantitative data. The discrimination rule is derived from a training data set made up from the exhaustive list of groups/categories of interest. When two groups are involved it is equivalent to finding a linear combination of the variables that would maximize the t-statistic for comparing the two groups.

Mahalanobis Distance: A weighted distance metric. The Mahalanobis distance between two points in one coordinate system is equivalent to the Euclidean distance between the two points in a transformed coordinate system where the transformation is derived from the variation in each dimension, the transformation from correlated to uncorrelated variables.

Ordnance Intensity Map: A plan map showing the probability distribution of ordnance intensity around *targets*.

Point deposition, point process: The process of adding ordnance to a *target*, or *targets* to a project *site*. Statistically, these are treated as one-dimensional, single point objects.

Receiver Operating Characteristics (ROC) curve: A graph of (percent or proportion false positive responses) vs. (percent or proportion of true positive responses).

Scale of correlation: A level of investigation in which the *point deposition* can be described by a statistical distribution. This report refers to three levels (*site*, *target*, and ordnance), all of which are controlled by independent *stochastic processes*.

Sector: A region or section of a project *site* considered to have homogeneous ordnance distribution under SiteStats/GridStats.

Site: An area encompassing an entire base or bombing range, usually containing more than one *target*.

Target: Any bombing target, impact area, detonation area, burial pit or similar feature that is likely to contain subsurface ordnance.

Target Intensity Map: A plan map showing the probability distribution of *targets* around a project *site*.

I Spatial Statistical Model and Optimal Survey Design for Rapid Geophysical Characterization of UXO Sites

II Performing Organization: Oak Ridge National Laboratory, Oak Ridge, TN

III Background

Sites that have been used for military weapons testing and training generally require remediation before they can be returned to public use. One of the hazards in such areas is the presence of unexploded ordnance (UXO) deposited over time. Partial characterization of UXO deposition at such sites is generally based on facility use records, or even geological features which suggest likely historical use. However, more precise and objective characterization requires physical screening of the area, generally based on indirect measurement technologies.

The existing statistical sampling approach, SiteStats/GridStats (SS/GS,QuantiTech, 1995a, 1995b), was developed in the early 1990s based on available survey technologies. The Department of Energy's Oak Ridge National Laboratory (ORNL) under the sponsorship of the U. S. Army Corps of Engineers Engineering and Support Center, Huntsville (USAESCH) recently reviewed the statistical methods in SS/GS (Ostrouchov et al., 1999). The SS/GS approach is based on limited statistical methods that do not use spatial information and do not provide confidence bounds on estimated contamination. Confidence bounds are currently provided by UXO Calculator (Barrett and Fanning, 1999), which also uses the assumption that contamination is uniformly distributed. As a result, subdivision of sites into sectors is artificially driven by an assumption that is required for confidence bound calculations and by the fact that UXO contamination is reported as a single number for a sector. At the same time, no statistically valid methodology is in place for testing the homogeneity of a sector. The homogeneity tests in SS/GS are not valid. As a result, sector homogeneity decisions are based on visual inspection of sampling results. In brief, the SS/GS procedure incorporates invalid statistical methods to characterize an inhomogeneous distribution of UXO. Further, without a valid test of sector homogeneity, confidence bounds produced by the UXO Calculator are invalid.

The fact that an extensive and complex methodology such as SiteStats/GridStats was developed and accepted, yet most of its statistical functions are invalid, should serve as an example that statistical methodology should be independently evaluated before being accepted. Experts who not only apply such methodology routinely, but also develop and publish within that field, should ideally perform such evaluations. For spatial statistical methods, this means statisticians who are involved in spatial statistics research. Complexity is no guarantee of validity.

Several geophysical methods for ordnance detection have been developed, and many of these have been deployed at the Badlands Bombing Range (BBR) in South Dakota. These include conventional ground-based magnetic and electromagnetic systems in addition to innovative systems; such as helicopter deployed magnetic and

electromagnetic systems (Doll et al., 2003) and the MTADS towed array system, which used both magnetometers and prototype electromagnetic systems (McDonald et al., 1996, 1998). Evaluation of airborne systems at BBR was funded by the Environmental Security Technology Certification Program (ESTCP) and has been highly successful (Fig. 1).

IV Objective

We address the site characterization problem in a comprehensive manner and provide specific solutions to several components of a comprehensive solution. Our objective in characterization is to provide maps that indicate areas of ordnance contamination (targets). A comprehensive solution means that we consider the entire span of activities from raw data acquisition through production of maps that indicate areas of contamination along with estimates of uncertainty. For the most part we do not differentiate between intact ordnance (UXO) and ordnance fragments as both indicate a possible target area location.

V Technical Approach

We begin by developing a conceptual site model that considers the process of site contamination, how we observe the contamination, and the resulting correlation structure. Based on this structure, on the physical nature of the contamination, and on the nature of our data, we develop statistical models that closely represent the underlying physics. On the basis of this Conceptual Site Model, we develop a characterization process that spans the initial information gathering from the archive search report (ASR) through platform and sensor selection, sample design, geophysical sampling, and final delineation of contaminated areas. Specific procedures are presented for most components of this process, while a description of needed components is presented for the remainder.

VI Summary

In the next Section, VII Project Accomplishments, we describe our accomplishments from two points of view: an overall conceptual site model (a developer's point of view), and an operational description (a user's point of view). In Section VII-1 we describe the Conceptual Site Model that concentrates on the underlying physical processes, and theory and underpinnings of our general approach. Specific methods are described in detail in Section VII-2, along with any further necessary theory and links to the underpinnings of Section VII-1. In Section VII-2, we follow a sequence in which the methods would typically be applied.

VII Project Accomplishments

Major concepts and tools developed include:
- Three Scales of Correlation concept (Section 1.3.1)
- DOAM – Distributions On A Map – OIM/TIM concepts (Section 1.3.2)
- Gauss.target simulation/ASR tool (Sections 2.2 and 2.7.1)
- Path characteristics optimization
- Target-related/Target-unrelated anomaly discrimination using multivariate statistical techniques applied to various signal characteristics.
- DOAM estimation with geoRglm (Section 2.6)

1 Conceptual Site Model: The Physical and Statistical Reality

1.1 The Physical Process of Ordnance Deposition

Procedures for characterizing site contamination should be developed so as to take advantage of what is known about the physical process leading to the distribution of ordnance. Briefly, individual items of spent ordnance (e.g., intact shells or shrapnel) are not typically distributed uniformly throughout a use-area but are rather in spatial clusters around some centers of activity. Within such clusters, individual ordnance items or *objects* are located at physical sites that tend to be close to the cluster center, relative to the typical distance between clusters. The purpose of the characterization is to locate those sub-areas that contain objects, or a sufficient concentration of objects to merit attention.

Our view is that ordnance is deposited at a given site through a series of activities that occurred over the entire history of that site. ***Two stages apply to each ordnance activity. First, a decision is made on the location or locations of the activity. Then the activity is performed near the locations and leaves behind ordnance in a cluster pattern that is determined by the activity***. For example, the first stage corresponds to a commanding officer's decision to locate a specific target and decide on the approach and ammunition to be used in that exercise. The second stage is the execution of the exercise, when ordnance is deposited in a random pattern mostly near the target. Another example might be a cleanup exercise to bury spent munitions. First a decision is made on what will be cleaned up and the location of the burial pit, and then the cleanup proceeds, leaving behind the buried ordnance and perhaps other scattered ordnance. Similarly, a firing range is first located at a convenient site and then ordnance is scattered with each firing practice.

Although we may notice that ordnance occur at higher density near the center of a target, we cannot measure this density at any location in the manner that we can measure a chemical spill concentration or ore content. Rather, we can use the ordnance locations to estimate the density per unit area. ***A physical process where data are locations of events (an event being the presence of a piece of ordnance) is called a spatial point process*** in the statistical literature (Cressie, 1991, Stoyan et al., 1995,

Diggle, 2003, Diggle and Tawn, 1998). We discuss the statistical issues arising from this physical process in Section 1.3.2.

1.2 The Data and Data Acquisition: How we observe the process

The locations of ordnance are not directly observed (unless exposed at the surface) and instead are estimated from indirect measurement technologies. Geophysical sensor systems, including magnetometers and electromagnetic (EM) systems are most frequently selected for this purpose. These sensors can be deployed on man-portable, towed, or airborne platforms (Fig. 1). An example of such a data set from Stronghold Table in the Badlands Bombing Range is shown in Fig. 2. Figure 2 is an analytic signal map, which is derived from airborne magnetic data, acquired with magnetometers (sensors) deployed on a helicopter (platform), as described by Doll et al. (2001). The unprocessed measurements consist of 12-meter swaths of data gathered with an array of eight magnetometers mounted on the helicopter at 1.7m spacing.   The data from each magnetometer are stored on a console in the helicopter. The individual data lines are processed, then gridded to produce a magnetic map of the site. The analytic signal is computed from magnetic data by calculating the square root of the sum of the squares of the magnetic gradients in three orthogonal directions. Therefore, the map in Fig. 2 and all similar geophysical maps represent a composite of data acquired along profile lines, with the values between lines and measured points along lines derived by minimum curvature gridding. This gridding algorithm interpolates data by fitting a two-dimensional surface to the raw data in such a way that the curvature of the surface is minimized. It yields best results when the data are expected to vary smoothly between measurement points, as is usually the case with potential field data (Geosoft, 1997).

Fig. 1. Recently completed upgrade of the airborne magnetometer system that was deployed at the Badlands Bombing Range in 1999. Data are acquired in a 12 m swath with 1.7 m sensor spacing.

Fig. 2. Magnetic map (analytic signal) from a portion of Stronghold Table at BBR. 20m grid cells.

1.3 Correlation Structures and Spatial Statistical Models

1.3.1 Three Scales of Correlation

We noted in Section 1.1, that ordnance deposition is a two-stage process. The first two scales of correlation are associated with these two stages. The first stage, which concerns the placement of targets (or activities that deposit clusters of items of ordnance), has correlations that are inherent in human decisions to place such targets. These may be correlations with topography, vegetation, and with activities that previously took place at the site (for example, decisions regarding minimum target separation, or safety, training and logistical considerations). The second stage governs the distances between individual ordnance items within such a cluster. We can easily observe that ordnance density (the number of ordnance items per unit area) tends to be very similar at two locations that are very close and usually not so similar at locations farther apart. This spatial correlation can be used to provide better estimates of ordnance density. Note that these two correlation structures are independent, as ordnance targeting precision is largely independent of how targets are spaced.

There is yet a third kind of correlation that concerns the similarity of an electromagnetic (EM) or magnetic signal at two very close locations in response to a single piece of ordnance. This correlation is a function of the nature of the signals, and the instrumentation used to generate them.

Three independent scales of correlation must therefore be considered in characterization of a contaminated site via EM or magnetic geophysical instrumentation:

- **Ordnance scale** (single geophysical anomaly scale)
- **Target scale** (multiple ordnance objects), and
- **Site scale** (multiple targets).

Although data maps such as these may be extremely detailed, they are subject to some degree of error. For example, the data acquisition is guided by navigation systems, but the sampling paths are never in complete alignment with their intended locations. Furthermore, the sensors respond to non-ordnance metallic objects, naturally occurring features (rock and soils), and electromagnetic interference, so that in spite of a wide range of analysis tools, a level of uncertainty always remains about the nature of the anomalies. Typically, it is not feasible to acquire geophysical data over an entire site. As an option, such systems can operate along "swaths" or "paths" to sample within the area of interest.

ORDNANCE SCALE: Geophysical response

500m

25m

60 km

SITE SCALE: Target placement, topography, vegetation, ordnance activity decisions

TARGET SCALE: ordnance and fragment placement

Fig. 3. Three independent scales of spatial correlation that can be exploited in site characterization.

EM and magnetic response signal correlation is at the single ordnance scale, ordnance placement correlation is at the single target scale, and target placement correlation is at the site scale. These are illustrated in Fig. 3.

Data collected by EM and magnetic methods provide direct information only at the single ordnance object scale. As a result, most characterization methods to date have concentrated on producing individual ordnance dig lists and simple site-wide averages (SiteStats/GridStats, UXO calculator). Two data conversions must occur to access the two larger scales of correlation. For target scale correlation, positions of individual ordnance objects must be extracted from geophysical data. Similarly, for site scale correlations, individual target locations must be identified. We describe these conversions in Sections 2.5 and 2.7, respectively.

Although we are focused on mapping contamination boundaries of entire targets, an understanding of physical characteristics of individual ordnance items, the physics of projectile impact, and associated geophysical signatures are important considerations in selecting signatures that are associated with a target. In particular, we must differentiate between geological anomalies and ordnance-related items. We address this in Section 2.5.

1.3.2 Spatial Representation of Current Knowledge: The Maps

Whereas representation of an uncertain quantity on a chart or profile can be shown with error bars, representation on a plan-view map really requires a collection of maps. We develop the concept of distributions on a map (DOAM, pronounced "dome") to describe contamination levels together with probability distributions of uncertainty. Our DOAM framework is based on the two-stage nature of ordnance deposition that we discussed in

Section 1.1, and on the need to properly account for uncertainty in our estimates. We focus on estimating two sets of DOAM maps, where each set corresponds to one stage of the two-stage ordnance deposition process. Estimation of each set of maps can take advantage of the respective scale of correlation. The map sets are:

- **Ordnance Intensity Maps (OIM)** provide information on density of ordnance at the target scale (we provide estimation procedures and programs).
- **Target Intensity Maps (TIM)** provide information on density of targets at the site scale (we only describe the concept of these maps).

We define intensity as the expected number of ordnance items per unit area rather than the actual number of ordnance items per unit area. If the ordnance deposition process were repeated many times and we averaged the number of ordnance items in a given unit area over the repetitions, the average would equal the intensity for that unit area. In fact, ***it is through the introduction of the concept of intensity that we are able to estimate the expected number of ordnance objects per unit area (i.e., the intensity) in areas not surveyed.*** *Expectation* has a precise definition in statistics (see, Hogg & Craig 1970).

Maps are not developed for the geophysical anomaly scale, because this source of variability is associated with the measurement system rather than the physical deposition of ordnance objects. However, we point out that such spatial analysis can be used to rigorously address the issue of non-detection for that measurement system by introducing spatial probability bounds on instrument signal. Such analysis would take into account any non-alignment and altitude variation of sampling paths and produce a location-specific estimate of non-detection. We emphasize that non-detection is really a spatially varying quantity that depends on local geological conditions and local signal sampling patterns. This is feasible with the spatial statistical methods used in this report, but is beyond the scope of this work.

The OIM and TIM DOAM maps are three-dimensional representations that include uncertainty as the third dimension. ***Our DOAM concept maps provide not only the estimated quantities but also a complete probability representation of the quality of our estimates***. Specifically, for a given map spatial resolution, each individual grid has a third dimension that describes the estimated probability distribution of ordnance intensity at that spatial location. The distribution can be summarized by a few quantiles, say .01, .02, through .99. In which case, a map of 100x200 grids is actually a 100x200x99 representation. Our software uses a sample of 1,000 intensities for each grid representation. Other representations are possible. The advantage of the DOAM three-dimensional representation and the ability to estimate it, which we discuss in Sections 1.3.2 and 2.6, is that it provides a rich set of maps that can be tailored for various sampling, remediation, and confidence assessment criteria. Some examples of maps that the DOAM representation can produce include:

- A map of probability that intensity is over a threshold per acre (see example in Fig. 4)
- A map of probability that intensity is under a threshold per acre.
- A map of probability that intensity is between two thresholds per acre.
- A map of locations where intensity is under a threshold per acre with at least 95% (or any other %) confidence.

- A map of upper 90% (or any other %) probability bound on intensity (see example map in Fig. 5).
- A map of lower 90% (or any other %) probability bound on intensity.
- A map of locations where we are at least 90% (or any other %) certain that intensity is under a threshold per acre.
- A map of locations we are at least 90% (or any other %) certain that intensity is under a threshold per acre (clean) or over a threshold per acre (contaminated).
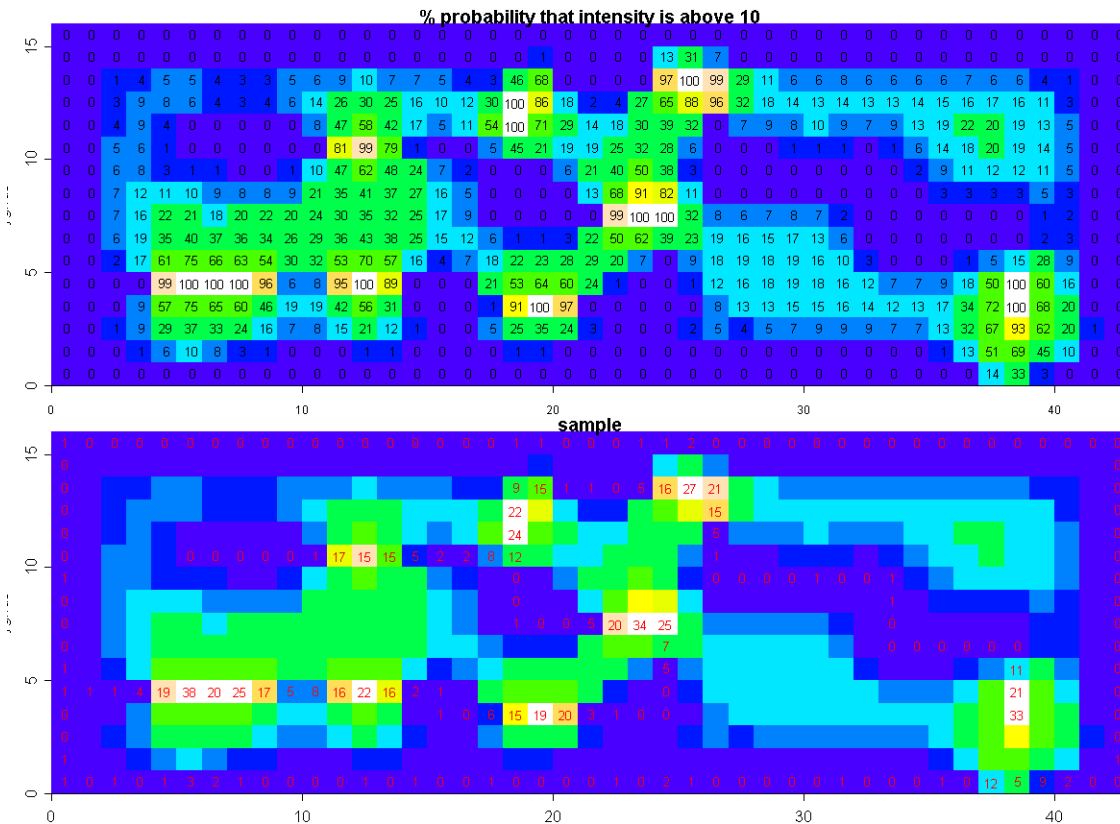
Fig. 4. OIM Map of probability that ordnance intensity is above 10 per grid (top) estimated from actual ordnance counts in a sample path (bottom). Color on both maps is keyed to probability on the top map.
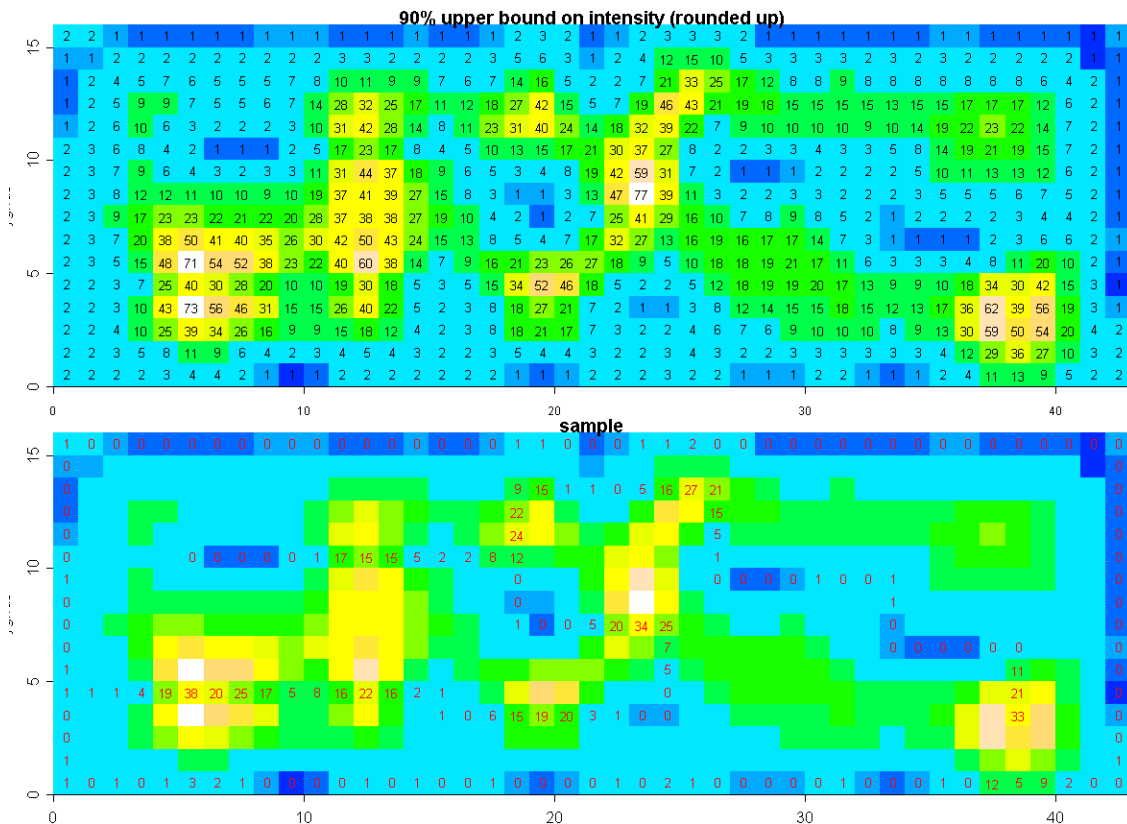
Fig. 5. OIM Map of 90% upper bound on ordnance intensity per grid (top) estimated from actual ordnance counts in a sample path (bottom). Color on both maps is keyed to upper bound on the top map.

Because we have the complete distribution estimate of the intensity at every grid of the map, we can map any combinations of intensity and probability that are necessary in a given remediation setting.

Maps can also be developed to drive further sampling by delineating areas that cannot be certified as clean (say 95 percent probability of being below some threshold) but have a substantial probability of being clean. This would be a map of areas with 50 to 95 percent probability of being below a threshold. Further sampling in these areas would narrow the confidence bands and allow more area to be declared clean.

The OIM and the TIM are developed initially from synthetic components that are based on a site conceptual model, which in turn is derived from the archive search report (ASR) and the associated topographic and vegetative cover maps, probable locations, types, and densities of UXO, land use maps, and other information, as we discuss in Section 2.2. The OIM and TIM are updated as survey data are acquired at the site, as prescribed by statistically based survey design procedures. Information from any additional area sampled is incorporated into an updated OIM and TIM, which in turn are used to produce specific maps to support the decision making process, and to guide survey design decisions if further surveying is required.

1.3.3 Statistical Models of Point Patterns
The deposition of ordnance, as we discussed in Section 1.1, occurs with greater frequency near a target than away from a target. This results in a non-uniform point pattern of ordnance locations. A process that deposits events (ordnance objects) scattered at non-uniform locations is described in statistics as an inhomogeneous Poisson point process (Cressie, 1993, Diggle, 2003). This means that ordnance objects are randomly placed according to an underlying inhomogeneous intensity (i.e., tend to be more densely deposited in some locations than in others).

One such inhomogeneous Poisson point process specification, the Neyman-Scott process (Neyman and Scott, 1958), mimics our physical reality of ordnance deposition very closely. In fact, Stoyan et al. (1995) note that Neyman and Scott (1972) use the process to model the geometry of bombing. Diggle (2003) describes it by three postulates (We add our ordnance deposition interpretation in parentheses.):

1. A spatial Poisson process generates parent events. (Activity locations such as targets are determined by a commanding officer.)

2. Each parent produces a random number of offspring independently and identically according to some probability distribution. (The number of items of ordnance used in one activity is independent of another activity.)

3. The positions of the offspring relative to their parents are independently and identically distributed according to a bivariate probability distribution. (The ordnance objects are scattered around the target(s) randomly according to some spatial distribution and independently of each other.)

This formulation leads naturally to simulation experiments to elicit properties of such a process and to compare various sampling strategies. We use such simulations to recommend path width and path spacing strategies for geophysical sampling in Section2.3.One of our central aims is to estimate the OIM from ordnance location data in surveyed areas. This means that we estimate ordnance intensity in the surveyed areas and predict ordnance intensity in areas not surveyed. Our estimation and prediction procedure provides uncertainty quantification for constructing the OIM representation discussed in Section 1.3.2. The estimation of maps generated by an inhomogeneous Poisson point process, such as the Neyman-Scott process is discussed in Section 2.6.

2 The Site Characterization Process and the Tools

This section describes the steps of the characterization process in the order that it would be applied in the field. Each subsection describes a step in this process, including any underlying statistical issues. The flow chart maps out the chronology of the sections as well as the characterization process.

Ideally, one should treat all three correlation scales in a single model so that full information from EM or magnetic signal samples can be carried through to target location estimates and contamination intensity estimates. Unfortunately, this presents us with intractable complexity. Good approximations exist if the problem is separated into ordnance location estimation and point process estimation. That is, we treat the smallest scale of correlation on its own and consider the output of this process as the input for

target and ordnance intensity estimation.
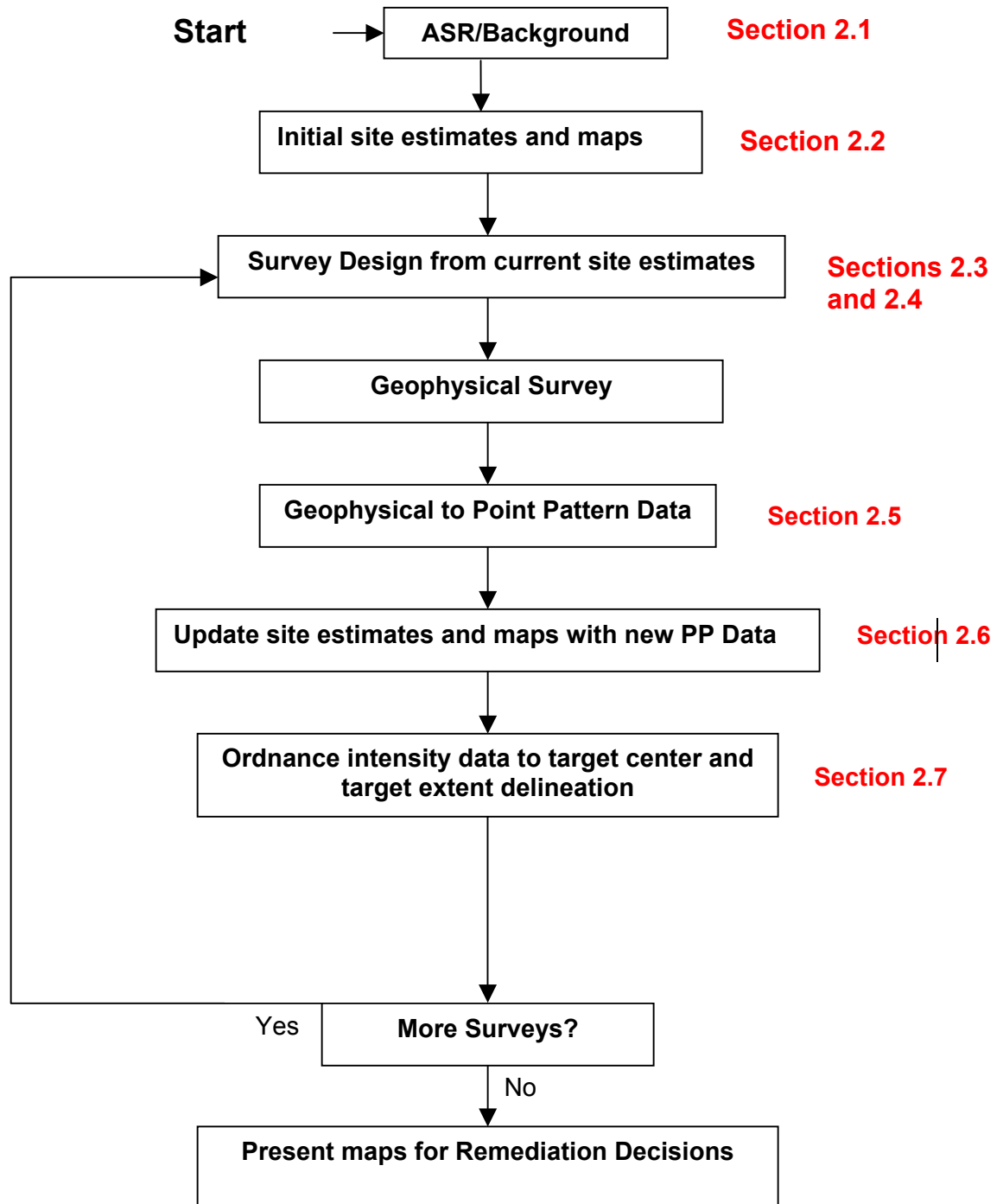
Fig. 6. Flowchart summarizing our method for designing and analyzing surveys for UXO.

2.1 ASR/Background

The starting point for a statistical UXO characterization is known as an Archive Search Report (ASR), usually conducted under the direction of the U.S. Army Corps of Engineers. They provide a description of the ASR process on their web site (USAESCH, 2003) as follows:

"The Corps of Engineers developed the ASR process as a cost effective means to determine the scope of potential hazardous material on former and active military installations. These potential hazardous materials include, but are not limited to, Ordnance and Explosive (OE), Chemical Warfare Material (CWM), and low-level radioactive material. The Center of Expertise and Design Center for OE at the U.S. Army Engineering and Support Center, Huntsville in association with the St. Louis and Rock Island Districts developed this process.

The ASR typically follows an initial small-scale study of a Formerly Used Defense Site (FUDS) called an Inventory Project Report (INPR). The local District of the Corps of Engineers produces this INPR. The INPR determines proof of past military ownership or use but they are limited in scope to data obtained from local sources. The core of the ASR process is the review and analysis of applicable textual records, maps and aerial photographs. This information is stored at numerous facilities including national, regional, state, and local archives and record holding facilities. The ASR team analyses the collected information for potential hazards. Other methods of information gathering may be beneficial such as interviews with veterans, former employees, and others associated with the sites. Many times interpretation of historic aerial photography greatly aids in identifying specific locations for potential hazards. The ASR team analyses the collected information for OE or CWM hazard potential. Interpretation of historic aerial photography greatly aids in identifying specific locations for this potential. An ASR site inspection follows after determining the areas to investigate. The inspection is limited in scope to a visual, non-intrusive inspection of the areas suspected as having a hazard potential. The ASR team follows a site safety and health plan prohibiting digging or handling of potential OE and CWM. Should any dangerous items be found, local law enforcement authorities are contacted to handle the immediate situation. Further actions depend on the circumstances.

The final ASR contains:
- *A brief history of the site*
- *Description and characteristics of the immediate surrounding area*
- *A review of related site investigations*
- *An aerial photography and map analysis of the site*
- *Real estate information, past and present*
- *Findings of the site inspection*
- *Description of the OE and/or CWM identified with the site*
- *Copies of pertinent documents gathered during the archives search"*

The ASR provides information for all three scales: ordnance type information for the ordnance scale, activity type for the target scale, and activity locations for the site scale. Of course, this information is incomplete and sometimes incorrect so that data are collected to complete, verify, and correct this information in a statistically rigorous manner.
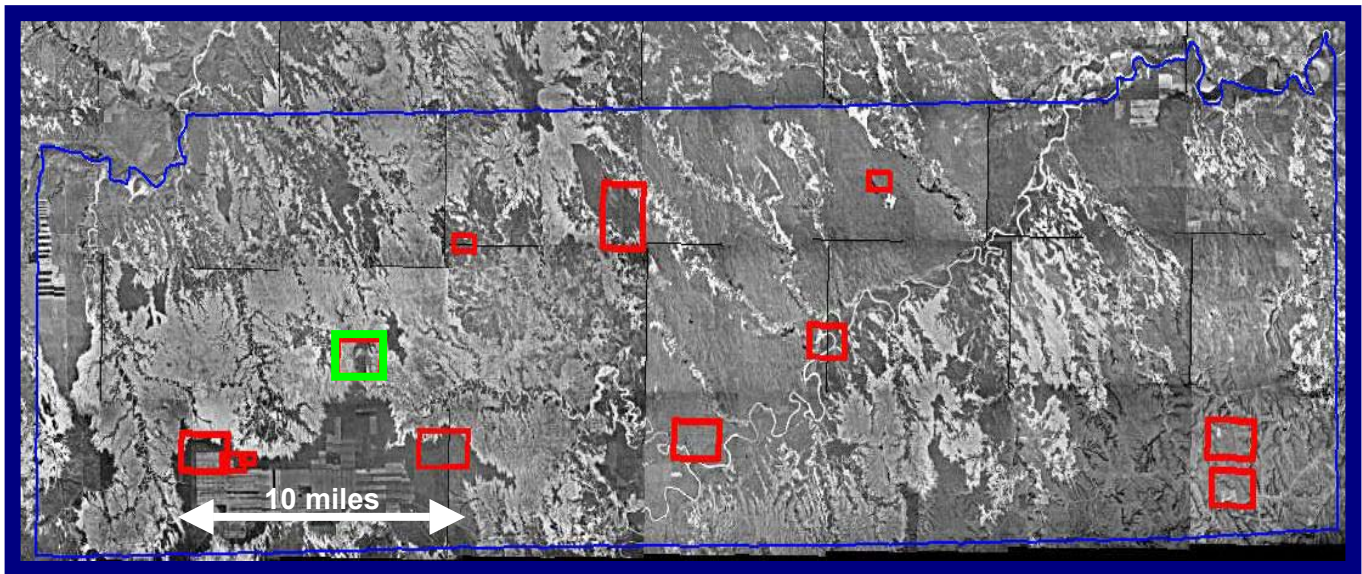
Fig. 7. Bombing Targets (shown in red) at the Badlands Bombing Range, South Dakota.
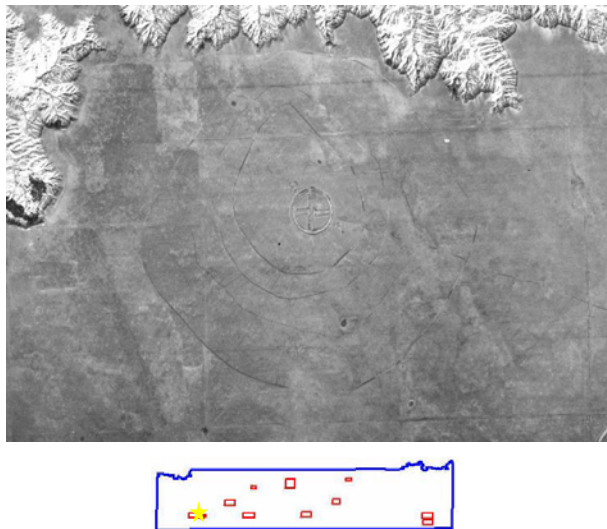
Fig. 8. Target 1 at BBR.
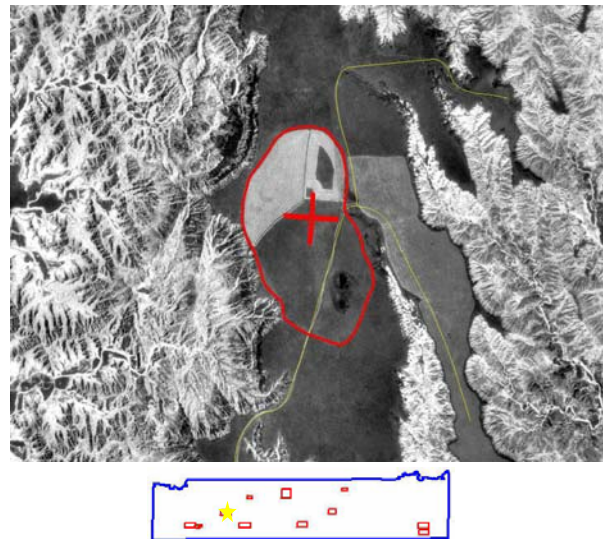
Fig. 9. Stronghold Table at BBR

The ASR for the Badlands Bombing Range (BBR) site (USAESCH, 1999) identifies 19 areas of concern within the 339,233-acre BBR site. This ASR was conducted following a 1997 NRL MTADS survey at two of the areas of concern.

2.2 Initial site estimates and maps

The final ASR, as described above, contains qualitative and quantitative information that can be used to construct an initial site map of target locations and rough estimates of potential ordnance contamination. In fact, the *"aerial photography and map analysis of the site"* is the step that must be supported with software that facilitates entry of target and contamination estimates into a GIS-based interface.

We have implemented a tool, "Gauss.target," in the R statistical software package (Ihaka and Gentleman, 1996) that allows the user to place targets on a map. All R functions associated with gauss.target are in the R-loadable file Rtarget on the accompanying CD-ROM. It is a simple example of how a much more complex and full-featured system would operate. The user is presented with a site map and graphical user interface (GUI) that allows the placement of suspected target locations on the map. Current implementation allows placement of targets with a Gaussian scatter of ordnance and computes site ordnance intensity on the basis of the user-placed targets and a specification of background intensity. An example of intensity that corresponds to the BBR map in Fig. 7, is given in Figure 10. Gauss.target will also scatter ordnance according to the specified intensity and output the ordnance locations from the map.

Gauss.target is only meant to illustrate the point that the ASR process should include software that allows electronic recording of quantitative information for use in later estimation and survey design. A full-featured tool would include a number of pre-specified activity types (bombing target activity, burial pit activity, artillery range, etc.) along with entry dialogs for parameters of the activity (dimensions of estimated ordnance activity, number of bombing exercises and number of ordnance per exercise, type of ordnance, etc.). An entry of a target along with its parameters would generate an upper and lower bound on intensity contribution of that target to the site. Such a tool would be somewhat like a number of current simulation games (Sim City, Age of Empires, etc.), where the user selects items to build on a map and the consequences of each selection and its parameters are recorded. The tool must also include features for
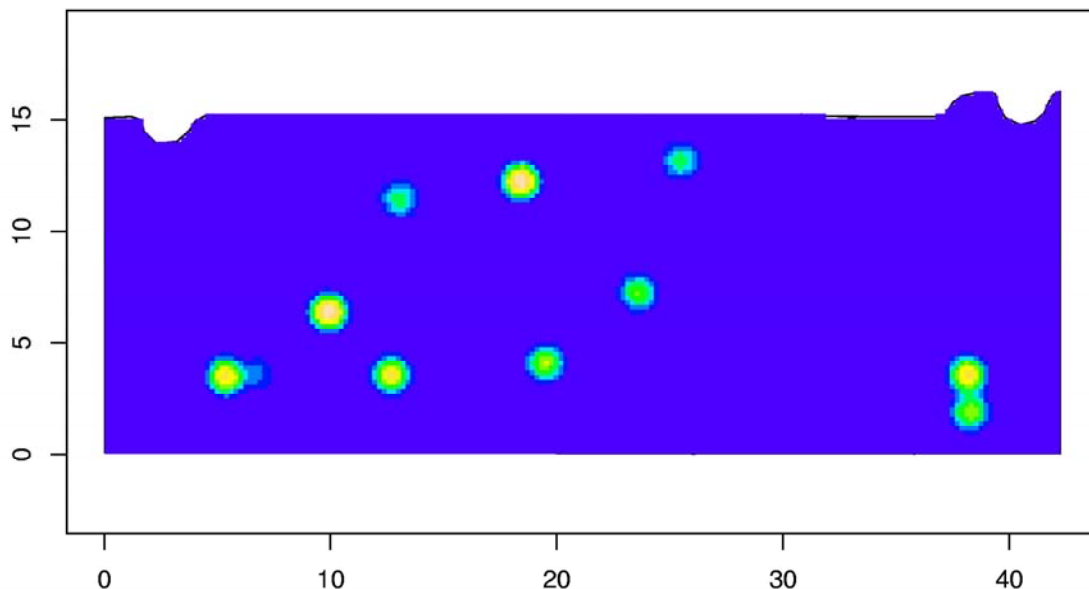
Fig. 10. Representative Initial OIM for BBR.

developing new types of targets. Clearly, this is a large project in itself, but it would provide appropriate quantitative inputs to the sample design and estimation process that follows.

Note that this tool would follow the two-stage nature to the assumed ordnance deposition process described in Section 1.1, where first a target is located and then its associated ordnance intensity is added to the site ordnance intensity. The output of this tool would be a three dimensional map, like the OIM, where the estimated intensity is provided at a specified spatial resolution, along with a measure of uncertainty in the third dimension. This measure of uncertainty can be simple upper and lower bounds computed by adding individual bounds provided with each entered target and some expert-elicited background estimate or possibly more quantiles that could be generated from distributional assumptions on each target type. In the simplest case of upper and lower bounds, the initial OIM would contain three intensity entries for each grid: the upper bound, the expected intensity, and the lower bound. The upper and lower bounds can be considered as the 1st and 99th percentiles.

We highly recommend that SERDP consider building such a tool that would provide ASR inputs to the methodology developed in this and similar projects. The activities recently developed under the names SimRange and Visual Sample Plan (VSP) are steps in the same direction as Gauss.target.

2.3 Survey Geometry Design

The sampling ideas we describe here follow directly from generic two-stage point-process models, such as the Neyman-Scott model, referenced above. Where the clusters of ordnance objects associated with each target are small relative to typical intra-cluster distances, much of the contiguous area of interest is UXO-free. Without precise prior information on where the targets are located, single-stage sampling plans will generally detect relatively few individual objects per linear unit of sampling path, since much of the path will be far from the actual clusters. In this situation, a two-stage sampling plan in which the goal of the first stage is cluster detection, and the goal of the second stage is the location of individual objects within each target-cluster, may be more effective.

Following this idea, we have focused our work on first-stage sampling plans for cluster detection. The prior information required includes two components. The first of these is a target-scale intensity function, which can be taken to be equal or proportional to the expected intensity of the initial OIM. The second is a set of conditional distributions of the number and spatial scatter of objects about a target at each potential location. Careful specification of this information would depend on a number of factors including the type of munitions used and spatial extent of range, much of which may also be derived from the ASR. Given these probability models for target intensity and within-target scatter, the probability that a given sample path fails to detect a randomly placed target can be calculated. By ``fails to detect a ... target,'' we mean fails to detect any object associated with the target. (Again, the idea is that if any event in a target-cluster is found, the second stage of sampling will be used to find the other objects in that cluster.) This failure probability can be written as a double integral involving events associated with any given potential target. The probability can be used as a

performance measure itself (i.e., the sample path might be more relevant for multiple, independently placed targets).Within this context, sampling plans made up of linear transects, meandering paths, or grids of relatively small disjoint sampled areas can be evaluated and compared. In its most obvious form, the integration defining the failure probability is not convenient for sample path selection because the sample path defines the region of integration for the inner integral; this requires that the entire calculation be performed for each possible path. Substituting a linear approximation for one of the factors of the integrand yields an approximation of this quantity as a single two-dimensional integral over the sample path. After some one-time "overhead" calculations, this approach greatly increases the speed with which each path can be evaluated (or alternatively, the number of possible paths that can be compared).

2.3.1 Model: Targets, Objects, and Paths

Procedures for designing sampling plans should be developed so as to take advantage of what is known about the physical process leading to the distribution of ordnance. Within clusters of ordnance, individual ordnance items or *objects* are located at physical sites that tend to be close to the cluster center, relative to the typical distance between clusters. The purpose of the survey is to locate those sub-areas that contain objects, or a sufficient concentration of objects to merit attention. Partial or vague information may be available about the size and extent of clusters. This information is not sufficiently detailed or certain to be the sole basis for ordnance removal. However, it can and should be used as the basis for designing sampling paths, the data that *can* be used as the basis for effective characterization.

The survey data can be collected by moving a detection device (for example, a magnetometer) along a path near the ground, within the area of interest. As the detection device passes over an object within its path, it signals the presence of the object (with some probability of error), providing an approximate location. Fig. 11 contains a simplified schematic showing clusters of objects denoted by filled circles, corresponding to five targets, each denoted by an open circle, and a sampling path comprised of three longitudinal linear transects through the area of interest. The objects are circles of various sizes, reflecting variation in the physical size of the objects. Detection methods have varying levels of sensitivity; relatively more sensitive methods might detect objects in three of the five clusters depicted in the figure, while less sensitive methods might detect only the larger objects shown in two of the clusters. (Note that "sensitivity," as used here, is related to "false negative" errors – e.g., the possible failure to detect an item of ordnance in the sample path. In reality, "false positive" errors attributable to objects which are not items of ordnance, but which are detected by the sensor systems, must also be considered.)
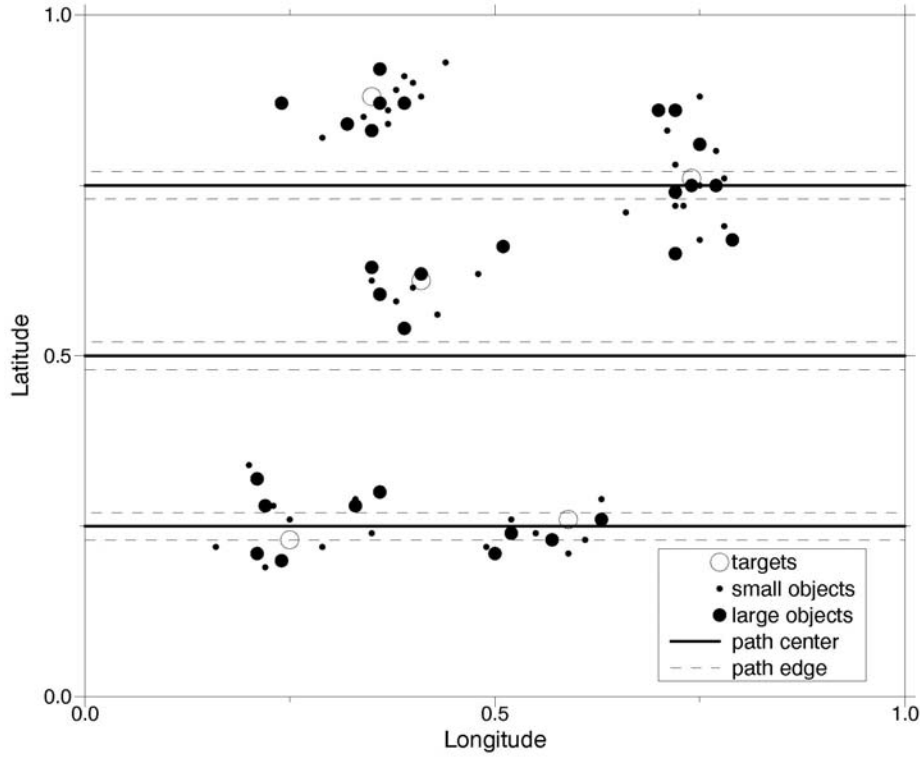
Fig. 11: Example of Targets, Objects and Paths. Diagram displays 5 targets, a cluster of large and small objects around each, and the centerline and edges of a sampling path comprised of 3 linear transects.

Because we are primarily interested in approximately locating the clusters, rather than precise definition of the extent of each, our primary goal will be to construct sampling paths which have the greatest chance of detecting at least one object in each cluster. Put another way, we wish to avoid a situation in which we fail to detect the existence of a cluster, and our approach is to construct sampling paths that minimize the probability that this happens.

Mathematical Formulation

The model we describe here is a more specific version of the Neyman-Scott model described above, which is a direct incorporation of the ideas just described. The spatial location of targets across the area of interest A follows a non-homogeneous Poisson process. For our purposes, we will consider the equivalent formulation in which the total number of targets, N is a Poisson variate with mean Λ. Individual locations of these targets t_1, t_2, ... t_N , are independently drawn from the distribution I(t). If a target occurs at t, the number of events resulting from that target, n_t, is a Poisson variate with mean μ_t. (Note: Strictly speaking, it might be more appropriate to use a truncated Poisson distribution omitting the possibility of $n_t=0$, since "targets" with no associated "events" are of no practical interest for our purposes. However, this distinction is not of practical importance when the mean number of events associated with each target is not small.) Each event associated with a target is determined as $e_{t,i} = t + \varepsilon_{t,i}$, i=1,2,...$n_t$, where $\varepsilon_{t,i}$ is a

17

bivariate random variable, independent of all other random variables, with distribution ϕ_t which may be specific to the target location.

The distribution intensity function for targets I(t), and the spread of the target-specific event distributions ϕ_t , are such that events generated by the same target are typically much closer together than events generated by different targets.

We assume that sampling (data collection) is executed by operating a sensor system along a "path" through A. Any particular path will be characterized as a relatively narrow "band" of width w centered about one or more continuous line or curve segments in A, denoted by p (as depicted in Fig.11). The specific process of "detecting" objects for a particular methodology involves the "physics" of background anomalies, how the signal is processed, and other details. These considerations also contribute to a careful and specific analysis of the origin of false-positive and false-negative errors. Here, we shall simply say that of the objects actually lying in the sampling path, a proportion s (for "sensitivity") is actually detected.

Under this model, the probability of failing to detect all of the objects associated with a target located at point t, using a sample path denoted as p, is approximately:

$$\text{Prob}(t,p) = exp\{ - \mu_t \{ s\, w\, /\, \sigma \}\, \phi(\, d_p(t)/\sigma\,)\} \qquad (1)$$

where:

  $s =$ proportion of objects that can be detected with the methodology used
  $w =$ width of sampling path (from edge to edge)
  $d_p(t) =$ smallest distance from t to the path center (measured perpendicular to path)
  $\phi =$ density function of the standard normal distribution

The approximation depends on (1.) the path-width, w, being small and (2.) the path being approximately linear both with respect to the scale of a cluster diameter, or about $4\text{-}6\sigma$. Note that this expression is conditional on the location of a target, t, and so does not involve the location of targets by the Poisson process described above. In Section 2.3.2 we consider the performance of sampling paths for such fixed (but generally unknown) values of t; Section 2.3.3 deals with unconditional probabilities and so involves $I(t)$.

As noted above, the parameter s (for "sensitivity") in this formulation allows for false-negative results from the sensor system. A model accommodating false-positive results can be developed along these lines by adding a relatively small "background" term, not associated with the spatial target clustering process, to the probability of detecting an "event" at any location. This would add some complexity to the model and we have not developed the methodology along those lines here, but such refinements are certainly possible provided data reflecting the density of detectable "geophysical clutter" can be obtained for a site.

2.3.2 Effects of Detection Methodology and Path Geometry

Characteristics of the detection methodology being used, and the sampling path followed, each have an impact on the probability of success of the screening operation. Other things being equal, one would expect to see improvements in performance— i.e., a decrease in the probability of overlooking a cluster—with increases in the sensitivity of the instrument, or in the width of the "swath" covered by the instrument as it is moved along the path. Likewise, increases in the length of the sampling path (and so total area sampled) or in the "uniformity" with which the path covers the area of interest should ordinarily correspond to improved expected performance. Hence, one set of specifications may be used to offset the effects of operational constraints on the other. For example, suppose an area to be explored consists of uneven or wooded terrain, so that airborne detectors must be used at a greater altitude than would be desired. This might ordinarily be expected to lead to a decrease in the instrument sensitivity that can be expected, and may make it especially difficult to survey certain sub-areas of the domain. Alone, either of these effects would lead to lower expected performance. However, in some cases it may be possible to compensate for this by increasing the length of the sampling path, and/or by adding width to the sensor array (e.g., widening the "path" covered).

Figure 12 shows how instrument and path characteristics interact in their effect on performance. In this graph, M and G are two summary measures of performance due to the sampling methodology (larger values of M are better) and path geometry (smaller values of G are better). The figure shows that curves of equal performance, indicating that smaller values of M can be, at least to some extent, offset by smaller values of G, and vice versa.
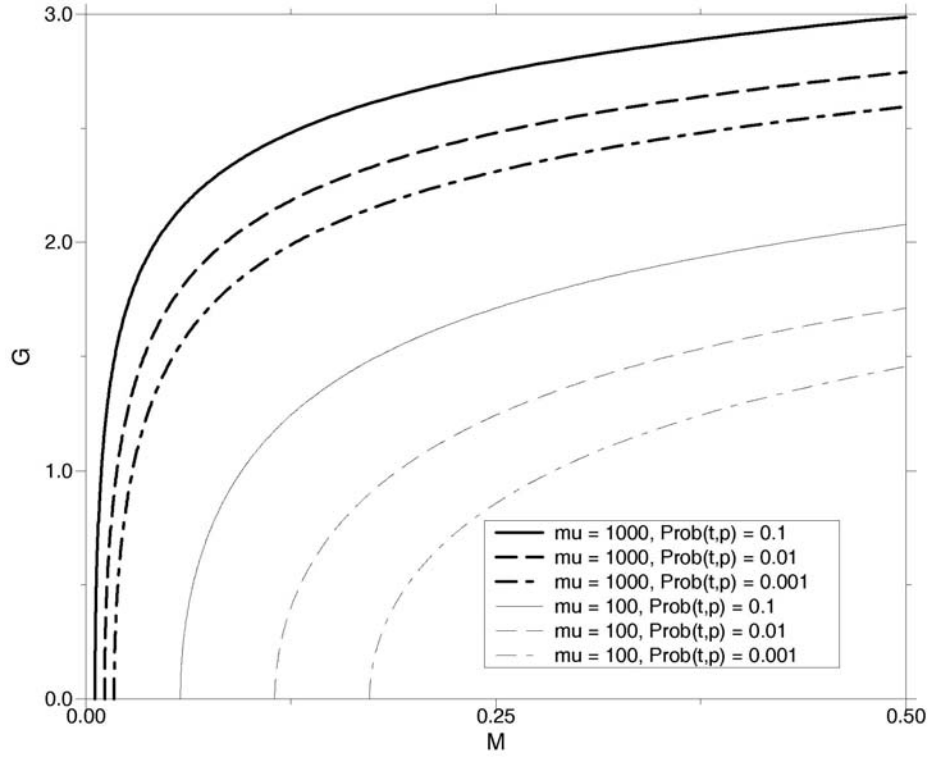
Fig. 12: Joint Effects of M and G. Plotted curves identify the range of (M,G) values that lead to three specified probabilities of detection, for two different expected cluster sizes.

Mathematical Formulation

Equation (1) may be rewritten in a form that more clearly summarizes the influence of detection methodology and sampling plan characteristics, as:

$$\text{Prob}(t,p) = \exp\{-\mu\, M\, \phi(\,G\,)\} \qquad\qquad (2)$$

where:

$M = sw\, /\, \sigma$
$G = d_p(t)\, /\, \sigma$

The "consolidated" parameters in this form of the expression are the unitless quantities M related to measurement technology characteristics, and G related to the geometry of the sampling path used. μ, the mean number of objects per cluster, remains and can be thought of as an index reflecting the overall difficulty of the screening problem, since smaller/larger values of μ represent clusters which are generally harder/easier to find, other things being equal. (We have dropped the subscript t on μ for the time being, as we are not concerned here with how it may vary across the area of interest.) The implication of this form is that, for a given path geometry, detection methodologies can be ranked for effectiveness by their corresponding values of M, where larger values of M

20

are preferred. Lower values of detection sensitivity, s, can be offset by wider paths, i.e., larger values of w, and vice versa. Detection methodologies with the same product sw will be equally effective, and both will be relatively more or less effective as the clusters are smaller or larger (as represented by the value of σ), respectively. Likewise for equivalent technologies of a given value of M, sampling paths for which G (for path geometry) is relatively small result in a relatively larger probability of detecting objects associated with a target placed at t. Paths with equal G are equivalent for detecting such a target, and all such paths are relatively more or less effective as the clusters are larger or smaller (as represented by the value of σ), respectively.

This expression is used as the basis for Fig. 12, which displays equivalent combinations of M and G values, for μ (average number of objects per cluster) values of 1000 and 100. So, for example, a detection methodology that is characterized by M of about 0.25 for a particular application is capable of detecting target-clusters of normalized distances of about 0.8, 1.2, and 1.6 from the path, with probabilities 0.999, 0.99, and 0.9, respectively, when the average cluster-size is 100 objects.

2.3.3 Two Popular Path Geometries

In the discussion above, the effectiveness of a particular sampling path for detecting a cluster centered at any point *t* is examined. In the formulae given, this is reflected as the factor $d_p(t)$, the shortest distance from the path to the cluster center. In this section, we examine this aspect of path geometry alone for two simple survey patterns of interest.

Suppose for this purpose we regard *t*, the location of a target, as being a random quantity, uniformly distributed across the physical region of interest. This corresponds to a uniform probability distribution *l(t)* in the mathematical description of the model in Section 2.3.1. With respect to this random distribution, d_p —now without specification of a particular *t* —is also a random variable, i.e., the distance between a selected sampling path *p* and a randomly chosen target location *t*. As noted above, relatively small path-to-target distances lead to relatively small probabilities of missing a cluster, other things being equal. Hence, particular sampling plans can be compared for overall effectiveness by examining their respective induced probability distributions of d_p.

Figure 13 displays two particular sampling plans of interest, each laid out in a square area of interest for this example. The solid lines represent path centers; path edges have been omitted in this figure because $d_p(t)$ is distance from the target to the center of the path. Panel A displays a sampling plan comprised of six parallel linear transects, while the plan shown in Panel B is made up of two perpendicular groups of three parallel transects each. Note that the total transect length is the same for these two plans. In each case, the transects are equally spaced, with the distance from a border transect to the area edge equal to half the common inter-transect distance. The open circles shown in each panel are potential target sites that are *most remote* from the path, that is, target sites that would be most difficult to detect with the displayed sampling plan. Note that there are more most-remote points shown for the plan comprised only of parallel transects; all points along lines mid-way between transects are most-remote for this plan, while only mid-points of the squares formed by perpendicular transects are most-remote for the sampling plan shown in Panel (b). However, the most remote points in Panel A are of distance 1/12 (relative to the side of the square area) from the path, while those in Panel B are of distance 1/6 from the path. For a general even number of linear

transects *m* (rather than 6), these values are 1/(2*m*) and 1/*m*, respectively. Hence for any (even) number of transects, the potential worst cases (i.e., most remote points) can be detected with greater reliability when all transects are parallel, rather than when they are arranged in two perpendicular groups. On the other hand, as will be shown in a later example (Section 2.3.5), even a few simple constraints can result in an optimal search pattern with crossing line paths.
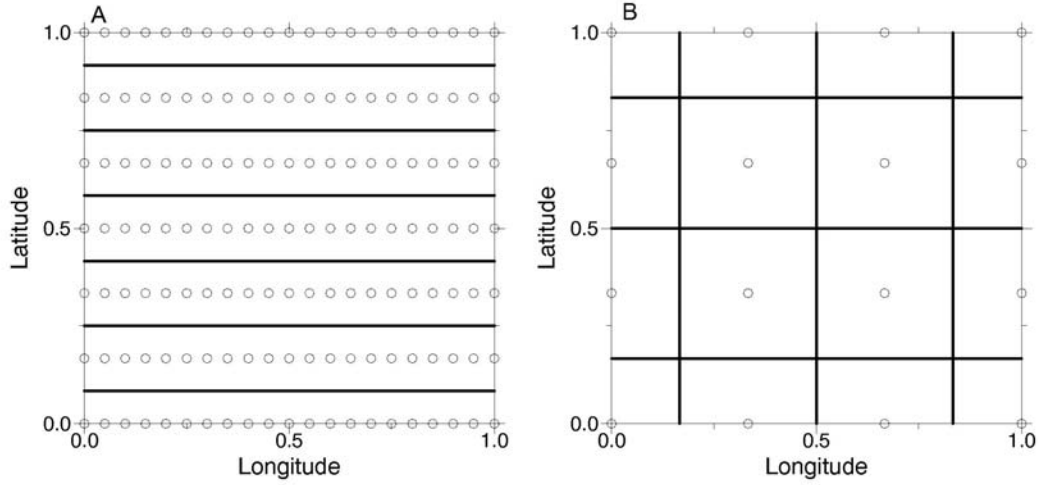
Fig. 13: Two Sampling Paths and Most Remote Points for Each. Sampling paths comprised of 6 linear transects are displayed as heavy lines, and the unsampled points farthest from the path are displayed as open circles.

Figure 14 displays the induced probability distributions of d_p, corresponding to the sampling paths displayed in Fig. 13. The upper extent of the horizontal axis for each graph is 1/6, showing that the points at this distance are most-remote for the sample path displayed in Fig. 13b. The means of these two distributions are 1/24 and 1/18, respectively. For a general even number of linear transects (rather than 6), the shapes of these distributions and ratios of the means remain the same, as the distributions are "compressed" along the horizontal axis. So, as with the distance to most-remote points, the distribution of distances to *randomly* selected points suggest that more effective screening is accomplished when all transects are parallel, rather than when they are divided into two perpendicular groups.
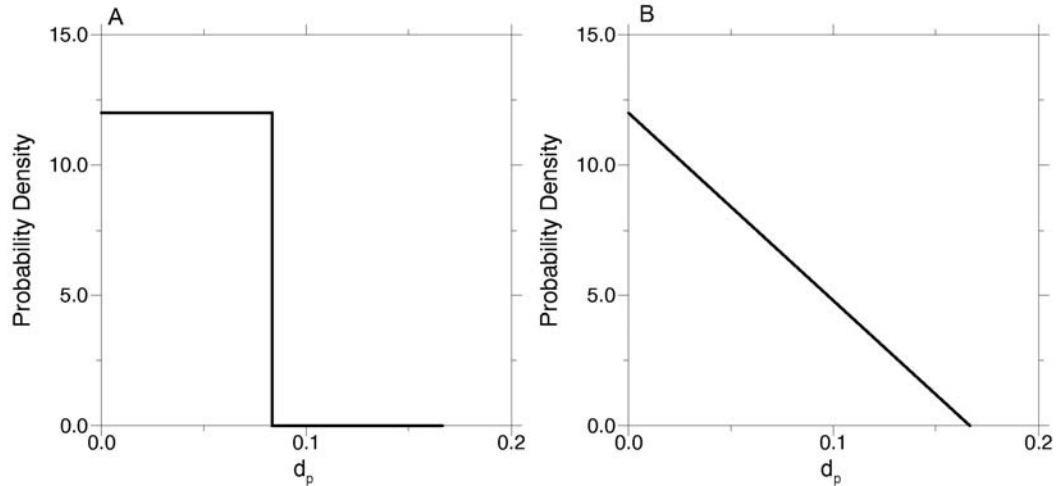
Fig. 14: The Probability Distributions of d_p for the sample paths shown in Figure 13.

Other path geometries of interest include S-shape or so-called "meandering" paths. Extreme S-shaped paths comprised primarily of parallel path segments connected with relatively tight turns at each end perform much like parallel transects. Under the model considered here, relatively good sampling paths generally:

- cover the area of interest uniformly, and
- contain few or no crossing segments.

The first of these properties is easy to see—paths that leave substantial sub-areas unsampled are especially poor at detecting clusters in those sub-areas. Paths that "cross themselves" frequently are generally wasteful for screening purposes since they commit multiple path segments to covering the same small sub-area.

2.3.4 Algorithm for Constructing Optimal Paths

The analysis presented in the previous section is intended to demonstrate general relationships between geometry of the sampling path and characteristics of the detection methodology. This is useful in showing the "broad" effects of screening characteristics, but may not be a sufficient guide for designing sampling patterns in specific situations. For a particular site, information may be available on the likely spatial distribution of UXO. Further, irregular terrain features may represent operational restrictions on practically useful sampling plans. Because the unique characteristics of a problem make it difficult to offer simple and general sampling rules for conducting a survey, we have written a computer algorithm for constructing sample plans which are "optimal" in a reasonable sense, given partial information about UXO spatial intensity, allowing for user control of which transect segments are considered for inclusion in the plan (i.e., are practical segments given what is known about the region).

Three types of information are required as input for the algorithm:

- information characterizing the intensity of targets across the area to be screened, along with the likely size and number of objects included in a potential cluster located at any point,
- information on the detection methodology to be used, in particular about the sensitivity of the method and the width of the path to be scanned along each transect,
- an explicit list of linear transect segments that would be acceptable (i.e., not operationally infeasible) for inclusion in a sample path.

Given this information, the algorithm can be used to construct a sample path of specified length, made up of segments from the list provided, which minimizes the probability of missing all objects associated with any target cluster. Hence, the intent is to construct a path that maximizes the chance of seeing at least one object from each cluster of UXO; it is understood that subsequent local-scale sampling will then be needed to determine the extent of each cluster.

The algorithm is a "local search," which implies that not every possible path that could be constructed is actually evaluated. A direct, complete comparison of all possible paths would be too computationally intensive for practical use. The approach taken relies on a "segment exchange" process, in which modifications of an existing path constructed by removing and adding individual segments, are evaluated, and changes that result in improvements are incorporated. The resulting path cannot be guaranteed to be the very best possible (in the sense of minimizing the probability of missing a target cluster) because not all possible paths are evaluated. However, the heuristics on which the search is based can usually be relied upon to yield paths that are near optimal in this sense.

2.3.5 Example

Figure 15 displays examples of the two functions needed as user input by the algorithm. The function in Figure 15a is the "target intensity" function, which specifies a relative probability of location for the targets located in the square area of interest (site-scale information. Figure 15a, can be thought of as a contour map of this function, indicating that targets are considered relatively more likely in the northeast corner and relatively less likely in the southwest corner. Figure 15b depicts a second kind of spatial information, the expected number of objects associated with a hypothetical target located at each point in the area (target-scale information). Again, the figure can be thought of as a contour map, in this case indicating that any clusters located in the eastern part of the area are thought to contain relatively more objects on average, and clusters located in the western part of the area are thought to contain relatively fewer objects on average. Note that the information related in Fig. 15b does not reflect the spatial probability of objects being located at various points (ordnance-scale), but the expected number of objects in a cluster *conditional* upon a target being located at each point in the area. Although some expert judgment is required in the construction of these functions for a given situation, much or all of the basic information needed should be included in the ASR.
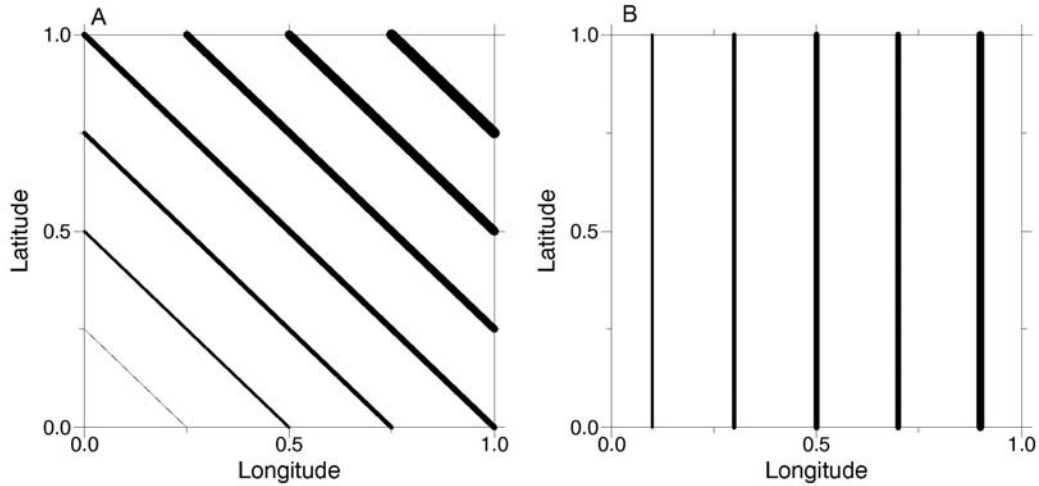
Fig. 15: Target Intensity Function, and Expected Cluster Size Function, for Example of Section 2.3.4. Panel A (left) displays contours of the target intensity function, ranging from 0.5 in the southwest corner to 1.5 in the northeast corner. Panel B (right) displays contours of the cluster size function.

The program also requires a list of path segments that can be considered as potential components in the search path. A collection of 88 path segments was specified for this demonstration. All segments are half the length of one side of the square region of interest. Half are laid out east-to-west and extend from the western boundary to the central meridian of the area or the central meridian to the eastern boundary. These are equally spaced north-to-south, with spacing equal to one-tenth the length of one side of the square region. The remaining half are similarly constructed, but run north-to-south, each one-half the length of one side of the square, and equally spaced from east-to-west across the region.

Figure 16, Panels A-D, display optimal paths constructed for this problem, comprised of 5, 10, 15, and 20 path segments, respectively. Calculations were performed for a detection system of perfect sensitivity (i.e., s =1), for situations in which the path-width is 0.0001 or 0.05 the length of one side of the square region; the resulting calculations were identical in this case. As can be seen in the figure, line transects are selected primarily along the east and north sides of the region in this case, reflecting the increased target incidence (Fig. 15a) and expected cluster sizes (Fig. 15b) in these regions.
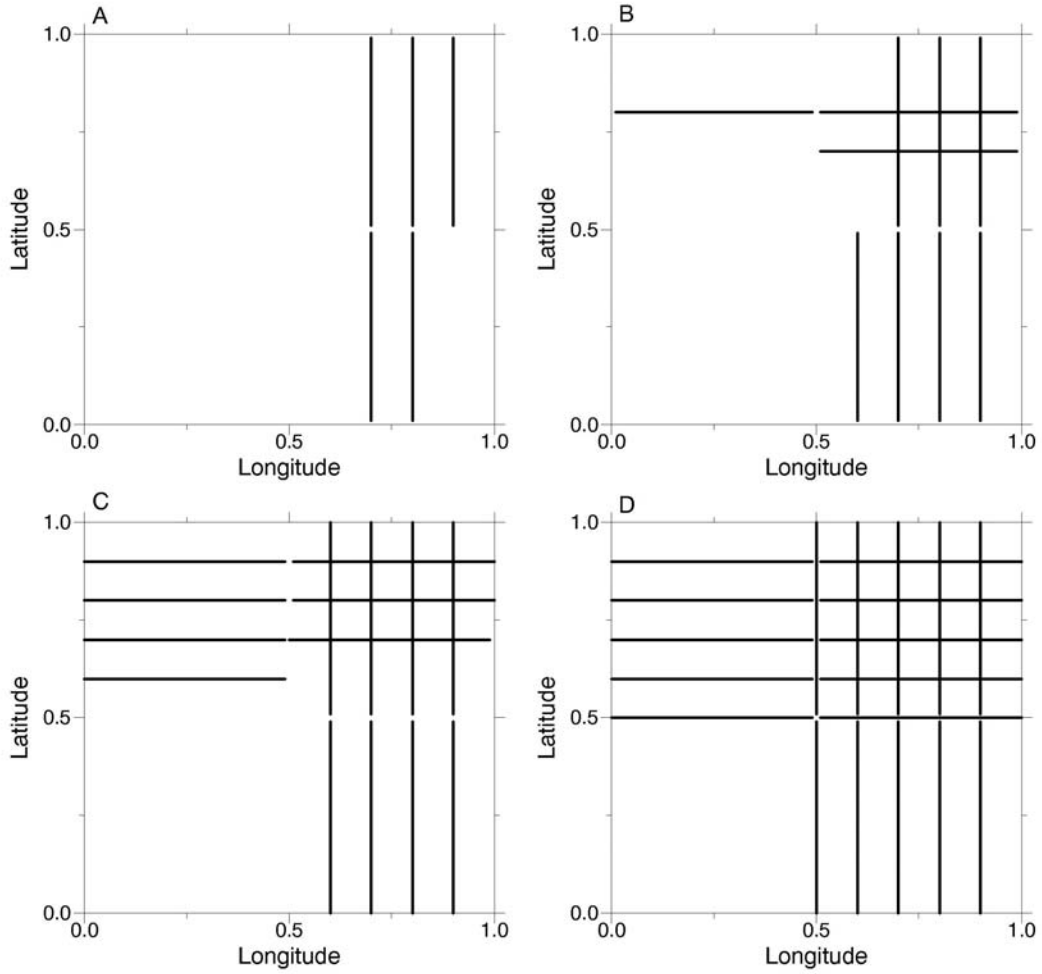
Fig. 16: Optimal Sampling Paths for the Example of Section 5. Paths comprised of 5, 10, 15, and 20 linear segments are displayed in Panels A, B, C, and D, respectively.

Mathematical details of the algorithm are described in the following subsection, and a listing of the S+ program used in the demonstration calculation is given on the enclosed CD-ROM..

Mathematical Formulation

If a target exists at t and exactly one event is associated with that target, the probability that the target is *not* detected by using sampling path p is:

$$1 - \int_{u \in p} \phi_t (u\text{-}t) \, du \;=\; 1 - \pi(t,p) \qquad (3)$$

where the integration region is the sampling path. Given the independent location of events within target-cluster, the probability of detecting none of the events associated with this cluster, conditional both upon a cluster being in this location and of producing n_t events, is:

$$[1 - \pi(t,p)]^{n_t} \qquad (4)$$

26

Removing the conditioning on the cluster size, this probability (still conditioned on target location) is:

$$\Sigma_n \, [1 - \pi(t,p)\,]^n \; \exp\{-\mu_t\} \, \mu_t^{\,n} \, / \, n!$$

Note that this may be rewritten as:

$$\exp\{-\mu_t \, \pi(t,p)\} \; \Sigma_n \, \exp\{-\mu_t(1-\pi(t,p))\}[\; \mu_t \, (1-\pi(t,p))]^n \, / \, n! = \exp\{-\mu_t \, \pi(t,p)\}$$

Finally, removing the conditioning on target location, we have that the probability of missing all events associated with a single, randomly placed target, is:

$$\text{Prob}(p) = \int_t \exp\{-\mu_t \, \pi(t,p)\} \, I(t) \; dt$$

Our goal will be to develop one or more algorithms which minimize Prob(p), or other related probabilities, with respect to the sampling plan, for a specified problem. However, Prob(p) as developed above is not a convenient form for path selection, because it requires the evaluation of a double-integral, with respect to each of t and u for any path to be considered. We next consider how this may be simplified to provide a more convenient form for our purposes. Recall that $\pi(t,p)$ is the probability that one unspecified event associated with a target at t will be detected using path p. If we assume this is a small probability, then we may express $\exp\{-\mu_t \, \pi(t,p)\}$ in the form of a Taylor series about $\pi(t,p)=0$ as:

$$\exp\{-\mu_t \, \pi(t,p)\} = 1 - \mu_t \, \pi(t,p) + \tfrac{1}{2} \, \mu_t^{\,2} \, \pi(t,p)^2 - \dots$$

Substituting into the above expression for Prob(p), we have:

$$\text{Prob}(p) = \int_t I(t) \; dt \; - \; \int_t \mu_t \, \pi(t,p) \, I(t) \; dt \; + \; \tfrac{1}{2} \int_t \mu_t^{\,2} \, \pi^2(t,p) \; dt \; - \; \dots$$

$$= 1 \; - \; \int_t \mu_t \int_{u \in p} \phi_t \, (u\text{-}t) \; du \; I(t) \; dt \; + \; \tfrac{1}{2} \int_t \mu_t^{\,2} \, (\int_{u \in p} \phi_t \, (u\text{-}t) \; du \,)^2 \, I(t) \; dt \; - \dots$$

$$= 1 \; - \; \int_{u \in p} \int_t \mu_t \, \phi_t \, (u\text{-}t) \, I(t) \; dt \; du \; + \; \tfrac{1}{2} \int_{u \in p} \int_{u' \in p} \int_t \mu_t^{\,2} \, \phi_t \, (u\text{-}t) \, \phi_t(u'\text{-}t) \, I(t) \; dt \; du' \; du \; - \; \dots$$

$$= 1 \; - \; \int_{u \in p} \Phi_1(u) \; du \; + \; \tfrac{1}{2} \int_{u \in p} \int_{u' \in p} \Phi_2(u,u') \; du' \; du \; - \; \dots$$

$$= 1 \; - \; \Psi_1(p) \; + \; \tfrac{1}{2} \, \Psi_2(p) \; - \; \dots$$

Hence paths which minimize the first-order approximation of Prob(p) are those for which the integral of Φ_1 over the path area are *minimized*. This is intuitive, because it favors paths that maximize the prior probability of detecting an event. However paths which are "optimal" by this criterion would occupy only the region in which Φ_1 is largest, avoiding regions in which the probability of detecting a target is even slightly less. Minimizing the second-order approximation of Prob(p) penalizes paths which cover the same area too many times since Φ_2 is relatively larger for such paths. Incorporating higher-order terms would yield greater accuracy, but require more intensive computing. Here, we shall focus on the second-order approximation of Prob(p),

$$\text{Prob}_2(p) \; = \; 1 \; - \; \Psi_1(p) \; + \; \tfrac{1}{2} \, \Psi_2(p)$$

Linear Path Segments

Substitution of $\text{Prob}_2(p)$ for $\text{Prob}(p)$ simplifies the numerical problem of finding optimal paths for target detection. Here we shall simplify the calculation further by restricting attention to paths which:

- have constant width, w, and
- are comprised of linear segments, each segment being all points in A within distance $\frac{1}{2}\, w$ of a specified line segment.

Let the set of line segments specifying a path be denoted by $L = \{\, l_1, l_2, \dots l_s \,\}$. For any single line segment l_i, the contribution to Ψ_1 may be approximated (for small w) by a one-dimensional integral along l_i:

$$\Psi_{1,i} \;=\; w \int_{u \in li} \Phi_1(u)\; du$$

Similarly, for sections of the path corresponding to segments l_i and l_j, the contribution to Ψ_2 may be approximated by a double line integral:

$$\Psi_{2,i,j} \;=\; w^2 \int_{u \in li} \int_{u' \in lj} \Phi_2(u,u')\; du'\; du$$

Hence we may approximate $\text{Prob}_2(p)$, for small w, by:

$$\Psi_{2,i,j} \approx 1 \;-\; \Sigma_{i=1}{}^{s}\, \Psi_{1,i} \;+\; \tfrac{1}{2}\, \Sigma_{i=1}{}^{s}\, \Sigma_{j=1}{}^{s}\, \Psi_{2,i,j}$$

Segment-Exchange Algorithm

- Given specified A, I(t), and ϕ_t,
- Given a specified list of "candidate segments", $\Lambda = \{l_1, l_2, \dots, l_t\}$,
- Given a specified number of segments to be included in the path, r < t,
- Given a specified path width, w,
- For each segment l_i from Λ , calculate:
  $$\Psi_{1,i} = w \int_{u \in ll} \Phi_1(u)\; du \;=\; w \int_{u \in li} [\int_t \mu_t\, \phi_t(u\text{-}t)I(t)dt\,]\; du$$
- Evaluate the u-integral over a one-dimensional grid of values along l_i.
- For each grid-point on l_i, evaluate the t-integral over a rectangular two-dimensional grid.
- For each pair of segments l_i and l_j calculate:
  $$\Psi_{2,i,j} = w^2 \int_{u \in li} \int_{u' \in lj} \Phi_2(u,u')\; du'\; du \;=\; w^2 \int_{u \in li} \int_{u' \in lj} [\int_t \mu_t{}^2\, \phi_t\,(u\text{-}t)\, \phi_t\,(u'\text{-}t)I(t)dt\,]\; du'\; du$$
- Evaluate the double u-integral over a (crossed) pair of one-dimensional grids along l_i and l_j, respectively. For each pair of points, evaluate the t-integral over a rectangular two-dimensional grid.
- Select an initial path $L = \{\, l_1, l_2, \dots l_s\,\}$ by randomly picking r segments from Λ . Denote by $\Lambda - L$ the set of t-s segments *not* in the initial path. Evaluate the initial path by calculating:
  $$\text{Prob}_2(p) = 1 - \Sigma_i\, \Psi_{1,i} + \tfrac{1}{2}\, \Sigma_i \Sigma_j\, \Psi_{2,i,j,}$$
  where sums are over the segments in the path.
- Find the segment l_i in $\Lambda - L$ which would most reduce $P_2(a)$ if it were added to the path, i.e. for which:
  $$\Psi_{1,i} \;-\; \Sigma_{j,\; lj \,\in\, L}\, \Psi_{2,i,j} \;-\; \tfrac{1}{2}\; \Psi_{2,i,i}$$
  is greatest. Add this segment to L, and delete it from $\Lambda - L$.

- Find the segment $l_{i'}$ in L which would least increase $\text{Prob}_2(a)$ if it were deleted from the path, i.e., for which:

$$\Psi_{1,i'} - \Sigma_{j \neq i',lj \in L} \Psi_{2,i',j} - \tfrac{1}{2} \Psi_{2,i',i'}$$

is least. Delete this segment from L, and add it to $\Lambda - L$.
- Re-evaluate $\text{Prob}_2(p)$ for the new L. Repeat this process (e.g. exchange of segments between and $\Lambda - L$) until $\text{Prob}_2(p)$ is not reduced further. Then report the resulting (locally optimal) path L.

The Test Problem

- A = unit square.
- $l(t) = t_1 + t_2$, i.e. 0 at SW corner and 2 at NE corner.
- ϕ_t = centered (i.e. zero-mean) bivariate normal with $\sigma_i = 0.1$, $\rho = 0$, $i = 1, 2$.
- Λ = segments with endpoints:

(0.0,0.0)	(0.0,0.5)
(0.1,0.0)	(0.1,0.5)
(0.2,0.0)	(0.2,0.5)
...	...
(1.0,0.0)	(1.0,0.5)
(0.0,0.5)	(0.0,1.0)
(0.1,0.5)	(0.1,1.0)
(0.2,0.5)	(0.2,1.0)
...	...
(1.0,0.5)	(1.0,1.0)
(0.0,0.0)	(0.5,0.0)
(0.0,0.1)	(0.5,0.1)
(0.0,0.2)	(0.5,0.2)
...	...
(0.0,1.0)	(0.5,1.0)
(0.5,0.0)	(1.0,0.0)
(0.5,0.1)	(1.0,0.1)
(0.5,0.2)	(1.0,0.2)
...	...
(0.5,1.0)	(1.0,1.0)

- $s = 1$.
- $w = 0.0001$ and 0.05.

2.3.5 Comparison to Other Approaches

The statistical methods used in spatial sampling problems are generally classified as being either *design-based methods* or *model-based methods*. Each has characteristics that may be thought of as advantageous or disadvantageous relative to the other, depending upon the application, prior information, and questions to be answered. The methods described in this report are model-based, because their development proceeds from a stochastic model of the spatial distribution of UXO. Specific spatial models used in this report are the Neyman-Scott model (Section 1.3) and other formulations of the Cox model (Section 2.6). The statistical methods developed for predicting and

estimating intensities, target locations, et cetera, are constructed in reference to this stochastic model, and do not require spatially random sampling for valid results. Our emphasis has been to develop methods based on models that seem most physically realistic in the context of what is known about UXO deposition and spatial distribution.

The methodology developed by Bilisoly and McKenna (2003) is also model-based. The spatial model, in this case, is the stationary random field model from which "ordinary kriging" is developed. While spatial predictions can be made without assuming the specific form of the spatial model, the standard errors of these estimators are generally developed under the assumption that the responses at any collection of points follow a multivariate normal distribution. This may be a reasonable assumption for some longer length-scales, where general fluctuations in intensity are of greatest interest, rather than the point-locations of targets, or the spatial density of individual objects associated with a cluster. Under these conditions, kriging methodology is computationally simpler than the methods we have developed. However, at shorter length-scales, or when the physical clustering mechanism motivating the Neyman-Scott point-process model is dominant, kriging models are less reasonable representations of reality.

Design-based methods typically depend on relatively weak assumptions concerning the problem structure, and rely on random sampling (sometimes stratified) for statistical validity. The *Visual Sampling Plan* software (Gilbert et al., 2002, Hassig et al., 2002) is based on methodology of this type. The greatest advantage of design-based procedures may be that they do not require strong assumptions concerning the deposition pattern of the material being sampled. As a result, inferences drawn from data collected under these plans may sometimes be less susceptible to criticism than those obtained through model-based methods. However, because these methods typically do not incorporate spatial models for the process under study, they generally require more data than model-based methods. A further limitation of design-based methods is that they generally do not provide a useful framework for *predicting* intensities or concentrations of material except with areas which have been randomly sampled, i.e., the absence of an explicit spatial representation makes it difficult to predict via interpolation or extrapolation.

Both model-based and design-based methods *can* be useful tools in the characterization of UXO distribution, and in fact, may well be used together at different stages of sampling. As suggested above, the kriging model of Bilisoly and McKenna may have practical utility for modeling general trends in intensity over moderate or long length-scales, while limiting the computational effort required for design and analysis. The design-based techniques of Gilbert et al minimize reliance on spatial assumptions, and may be most useful when many spatially scattered samples can be quickly collected, and when only average levels of contamination within relatively large cells must be estimated. The methods described in this report are more computationally demanding, but are based on point-process models that may more faithfully represent the physical problem. In particular, the techniques to be described in Sections 2.6 and 2.7can be used as a "meta-method" to guide sampling where it is most needed over several length scales. Depending on the scale, operational constraints on sampling, et cetera, *VSP*, kriging, and methods based on point processes might *all* be used at different stages of a study and the results integrated through modeling based on physically realistic point process models.

2.4 Geophysical Sensor and Platform Selection and Performance

In Section 2.3.1, we have provided a mathematical framework for selecting optimal geophysical sensors and platforms for a survey. Equation 1 (p. 18) presents the probability of failing to detect all objects associated with a target located at point t, using a sample path denoted as p, as a function of four variables, s, w, $d_p(t)$, and ϕ. Two of these variables, s (the proportion of objects that can be detected with the methodology used) and w, (the width of the sampling path) can be used to quantify instrument and platform performance, in order to select the optimal tools for a particular survey. Moreover, it is essential to develop an understanding of geophysical anomalies at the ordnance scale in order to more accurately separate them from non-ordnance anomalies, and from these anomalies to construct point pattern distributions (Section 2.5), which are in turn used to make ordnance intensity maps at the target scale (OIM, Section 2.6).

2.4.1 Platform selection

Several platforms are available for UXO instruments, including man-portable, vehicle-towed, aircraft, or boats. Some platforms are inappropriate for certain sites, e.g., boats are inappropriate without a suitable body of water, and aircraft cannot fly at low altitudes where they are impeded by tall trees. The platform controls the minimum and maximum distances of the sensors relative to the object. In selecting a platform for a site, one must be concerned with several factors:

- Vegetative cover can eliminate one or more platforms from consideration. Tall trees can prevent airborne platforms from operating at sufficiently low altitude; dense forest can prevent operation of towed instruments and hinder performance of man-portable platforms. Swamps or dense underbrush may be unsuitable for man-portable or towed platforms.
- Rugged topography may restrict all platforms from consideration; moderate topography may only be suitable for man-portable instruments (Figure 17).
- The size of the UXO targets must also be considered. Airborne platforms operate at an altitude that is too great for detection of 20mm or other small ordnance objects. For intermediate sizes of UXO, it may be that instruments operating on one platform may be able to detect a larger proportion of UXO, although the cost per area may differ considerably.
- Noise levels vary among platforms, and can affect the detection capabilities of instrumentation deployed on that platform.
- At some sites, the presence of hazardous objects, chemicals, or explosives could disallow platforms that must interface with the ground surface.
- At some sites, regulations for the protection of sensitive species will not allow clearing of brush or disturbance of soils that are associated with ground-based platforms.
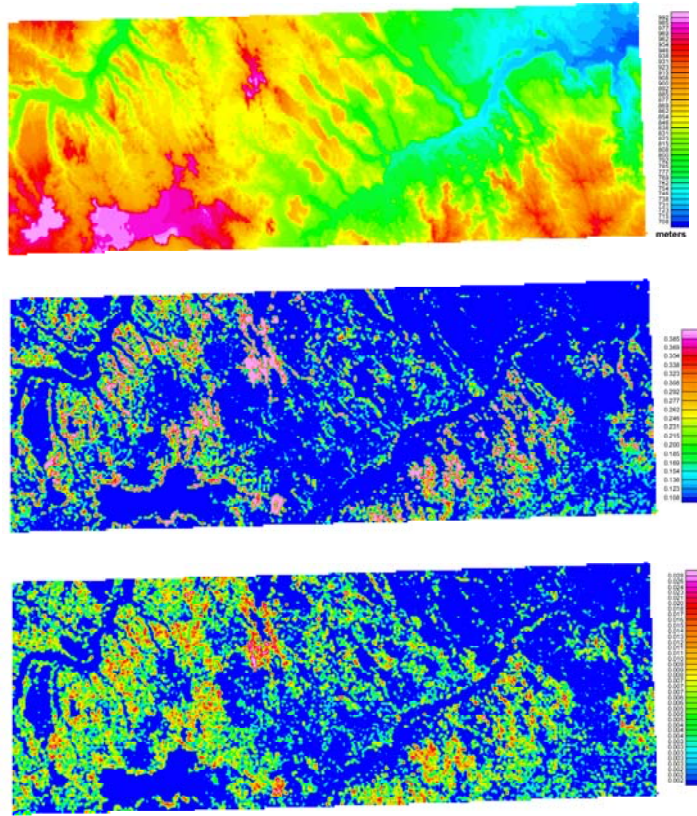
Fig. 17. a) Topographic map of the Badland Bombing Range, SD; b) first order gradient of the topography from a); and c) second order gradient of the topography, derived from b).

2.4.2 Instrument Selection

Most sensors used for UXO mapping are geophysical instruments, primarily magnetometers and electromagnetic induction instruments. Many of the electromagnetic instruments were originally designed for geologic mapping, although some have been adapted for detection of UXO and other metallic objects, and new systems designed specifically for UXO are in development. As they are passive instruments, magnetometers have generally not required design changes for UXO surveys. When operating at low sensor heights, it has been observed that electromagnetic instruments are usually more effective than magnetic surveys where the background geology causes significant interference. For instance, electromagnetic instruments are generally the instrument of choice in parts of Hawaii and the southwestern United States where basalts are predominant.
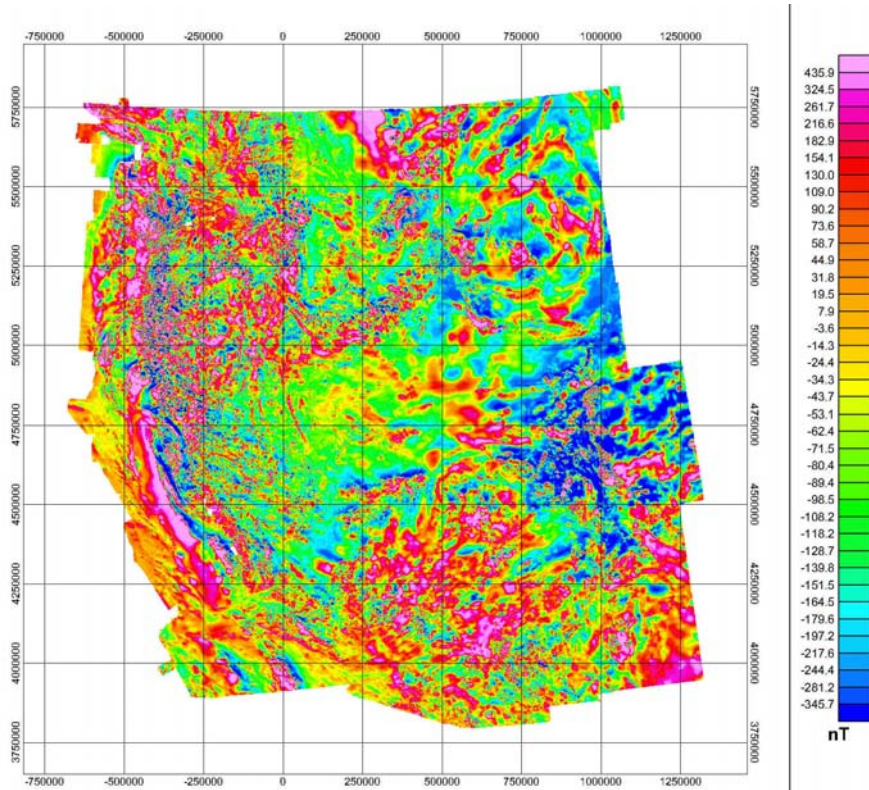
Fig. 18. Regional magnetic map of the western U.S.
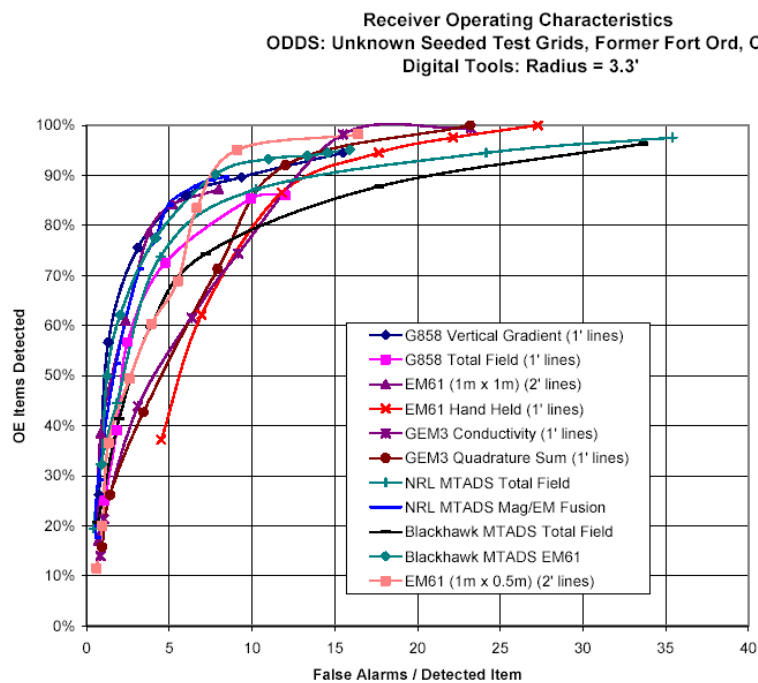
Fig. 19. Receiver operator curve from Asch et al., 2002.

33

Performance of sensors can be anticipated by reviewing regional geologic or magnetic maps, e.g., Fig. 18. Where the magnetic field due to geologic effects is large and variable over short wavelengths, it is quite likely that geology, associated with basalts or other mafic rock types, will interfere with magnetic systems, making electromagnetic sensors the preferred tool. A larger scale map than that shown in Fig.18, will probably be required in order to make this assessment. Selection of a sensor is most reliably determined on the basis of controlled tests, where the different alternatives can be judged in terms of their sensitivity to known targets. These may yield Receiver Operator Characteristic curves (ROC curves), as in Fig. 19.

ROC curves, such as the one shown in Fig. 19, provide an indication of sensor performance. They provide an indication of the ratio of false alarms to ordnance items, based on excavation of anomalies detected with a particular instrument or by a particular survey provider. Underlying dependencies to these curves include the range of types and depths of ordnance encountered, the accuracy of data positioning, the discrimination capabilities of the interpreter and the evaluator's differentiation between OE and non-OE items (intact-partial-pieces-frag). ROC curves can be used to select the optimal instrument, but do not provide a value for the parameter *s* in equation 1, which quantifies the proportion of objects that can be detected with the methodology used. This depends on how many anomalies are selected for excavation. The threshold that is selected in picking anomalies is directly related to the value of *s*. We suggest that the ROC curves be used for selection of the optimal instrument, and that the user assign a value to *s* based on the performance of the selected instrument at a local prove-out grid. If 80% of the UXO items of the desired type are detected in the test grid, using an anomaly threshold that would be used in the survey, then *s* should be set at 0.8.

Swath width, *w*, can be taken as about twice instrument height for single sensor instruments (1.5-3 times depending on the sensor, see Fig. 20). This will apply largely to ground-based instruments, which are operated at about 0.5 meters above ground level (AGL). For instruments that consist of arrays of sufficiently densely spaced sensors, the swath width can be estimated as the maximum separation of sensors plus twice the
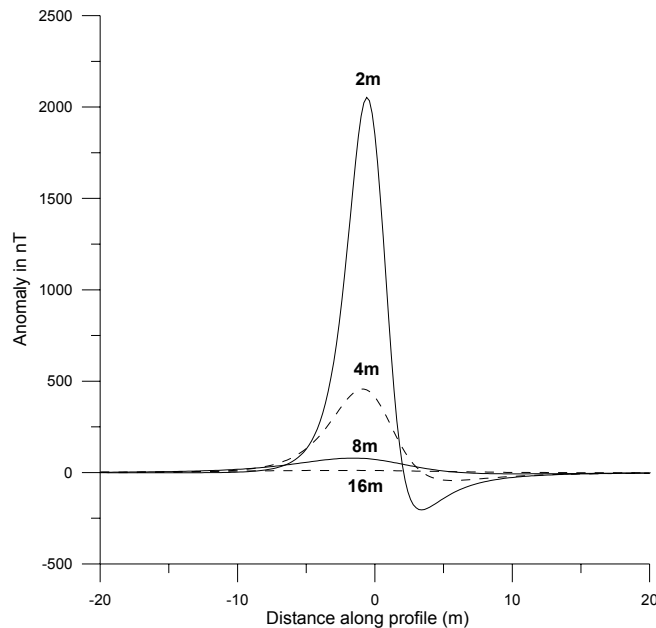
Fig. 20. Calculated response for a magnetic dipole for four selected sensor heights.

instrument height. The ORAGS-Arrowhead airborne magnetic system, for example, has a maximum sensor separation of 12m, and under favorable conditions operates at 1.5m AGL, leading to a swath width of 15m.

Practical considerations must guide the selection of platforms and sensors. Random paths cannot be programmed into navigation devices for airborne of towed sensor platforms. Actual paths will always differ somewhat from the programmed path locations. Vegetation or localized topographic features will often cause diversion from pre-programmed paths.

2.4.3 Geophysical Forward Modeling for Signature Prediction

The spatio-temporal characteristics of specific sensor signatures (e.g., magnetometers, time-domain electromagnetic induction systems and frequency-domain electromagnetic induction systems) of specific types of ordnance can be determined by geophysical forward modeling. Validated forward models exist for all the major classes of sensor systems used for UXO detection surveys. The concepts of two specific forward models for total field magnetics (TFM) and time-domain electromagnetic induction (TDEM) are illustrated in Fig. 21 (Butler et al., 2003). The models summarized in Fig. 21 have been extensively validated. The models also serve as the foundations for parametric inversion to obtain model parameters that best fit field survey data for localized, UXO-like target anomalies (Butler et al., 2003). Algorithms for discrimination of UXO-like from non-UXO-like targets are based on the model parameters recovered from inversion. Examples of model validations are shown in Fig. 22. The validation example for magnetometry compares forward model calculations with measured signatures over a 105-mm projectile at four azimuthal orientations (Fig. 22a). For TDEM, the validation example illustrates the use of the forward model in parametric inverse modeling and compares measured signatures for a 60-mm mortar and a 105-mm projectile with forward model calculations using parameters deduced from inversion of the measured data (Fig. 22b). The TDEM validation example illustrates capability to adequately model both the spatial signature and temporal evolution of the signature for two different types of ordnance.

Magnetic Forward Model

➤ **Approximate targets as a ferrous spheroids**

➤ **Express response in a multipole expansion**
   (McFee 1989; Altshuler, 1996;Butler et al. 1998)

   ➡ Anomalous, induced magnetic field: $\mathbf{b}^S \approx \mathbf{b}^{(2)} + \mathbf{b}^{(8)}$

- **Odd order terms zero due to axial symmetry**
- **Quadrupole term zero due to fore/aft symmetry**
- **Octupole term negligible for sensor-model distances > the major axis length**
- *Earth's field characterized by magnitude, inclination and declination*
- **Results in 8-parameter model vector**

   ➤ $m = [X,\ Y,\ Z,\ \phi,\ \theta,\ a,\ e,\ \mu_r]$

   Location   *Orientation*   *Size, Aspect Ratio*   *Relative Magnetic Permeability*

$r = |R| = |(X,Y,Z)|$

Prolate Spheroid

a. Concept of a prolate spheroid forward model (MAGMOD) for UXO magnetic signature prediction.

TDEM *Forward Model*

Express TDEM response as that of a pair of orthogonal dipoles at the center of an *axisymmetric* target,

$$\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t} = \frac{\partial \mathbf{B}_1(\mathbf{r},t)}{\partial t} + \frac{\partial \mathbf{B}_2(\mathbf{r},t)}{\partial t}$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are secondary magnetic fields due to the induced dipoles

$$\mathbf{m}_1(t) = L_1(t)(\hat{\mathbf{z}}' \cdot \mathbf{B}^p)\hat{\mathbf{z}}'$$
$$\mathbf{m}_2(t) = L_2(t)[(\hat{\mathbf{x}}' \cdot \mathbf{B}^P)\hat{\mathbf{x}}' + (\hat{\mathbf{y}}' \cdot \mathbf{B}^P)\hat{\mathbf{y}}']$$

and each dipole decays independently as

$$L_i(t) = k_i(t + \alpha_i)^{-\beta_i}e^{-t/\gamma_i}.$$

giving a 13-parameter model vector:

$$m = [X,\ Y,\ Z,\ \phi,\ \theta,\ k_1,\ \alpha_1,\ \beta_1,\ \gamma_1, k_2,\ \alpha_2,\ \beta_2,\ \gamma_2].$$

$\mathbf{B}^P$ $\mathbf{B}^P$

$m_2(t)$ $m_1(t)$ $m_2(t)$ $m_1(t)$

Plate-like   *Rod-like*
Axisymmetric Targets

b. Concept of an orthogonal dipole forward model for UXO TDEM signature prediction

Fig. 21. Forward models for UXO signature predictions

36

Magnetic Model Validation Example
Measured (NRL)

Azimuth = 0   Azimuth = 45   Azimuth = 90   Azimuth = 135   Azimuth = 180

Calculated

105 mm Artillery Projectile
Comparison of Measured Vs. Calculated Total Magnetic Field Anomaly
(Inclination = 0; Depth = 0.5 m; Sensor Height = 0.25 m)
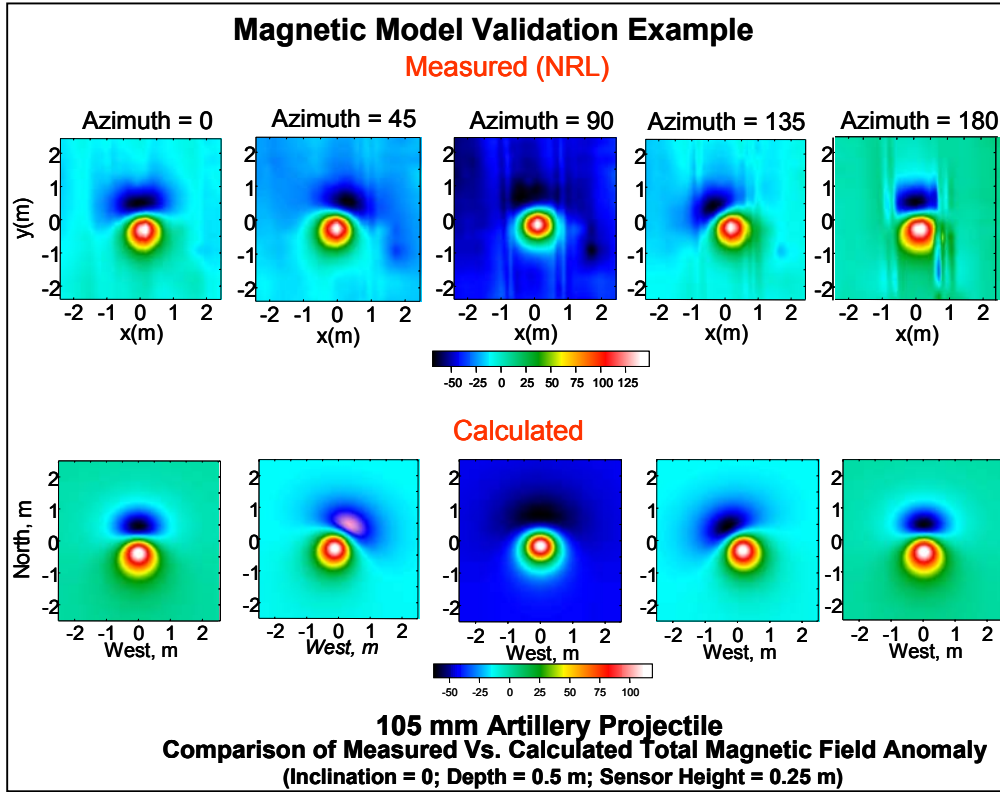
a. Forward magnetic modeling validation example for magnetometry

60-mm Mortar   105-mm Projectile
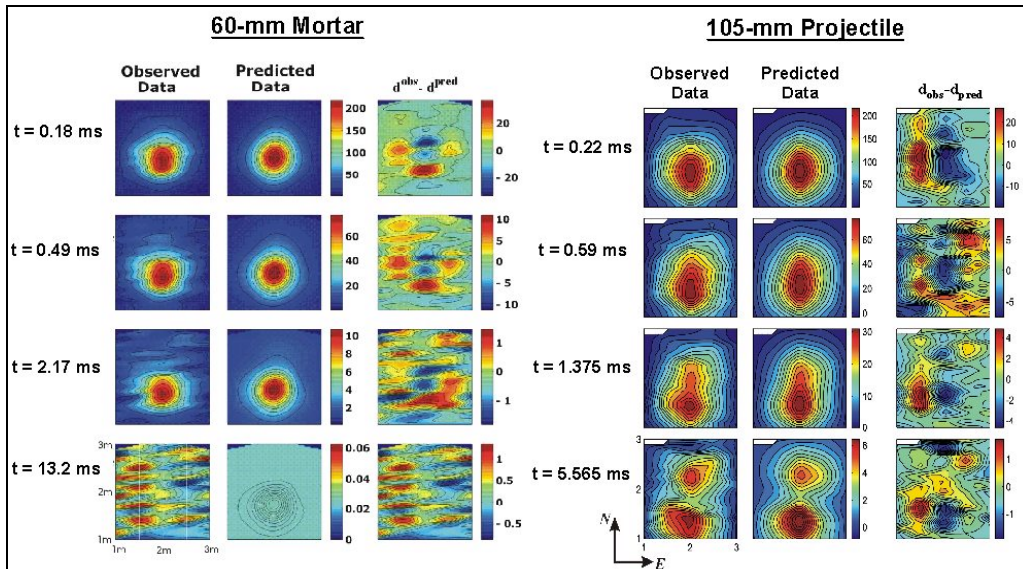
b. Parametric inversion and forward modeling validation examples for TDEM

Fig. 22. Validation examples for magnetic and TDEM modeling

2.4.4 Ordnance Types, Depths and Orientations

When the specific types of ordnance at a site are known, modeling may be used to represent the expected anomalies at representative depths and orientations. Representative noise values will be known for appropriate instruments, and this can be used as a threshold to select instrumentation for a site. Total magnetic field signature predictions have been developed for 105's, 155's, and M-38 practice bombs, and others can be developed. An example is provided in Fig. 23.   This figure shows the maximum positive and negative amplitudes for an M-38 practice bomb, the predominant type of ordnance found on the Badlands Bombing Range, as a function of sensor height above the bomb, for three different orientations of the bomb. These calculations were performed using the code MAGMOD to model the magnetic response of UXO (forward modeling; Butler et al. 1998; Butler et al. 2001). MAGMOD requires details of the ordnance item (length; diameter; distance below sensor, where distance = depth + sensor height; location within calculation grid; and orientation; see Fig. 21a). For magnetic modeling, the magnitude, inclination, and declination of the Earth's magnetic field at a location of interest is required. Several model UXO orientations should be calculated in order to assure that UXO will not remain undetected due to unfavorable orientations – the detection method must be able to detect those items as well as items that are favorably oriented.
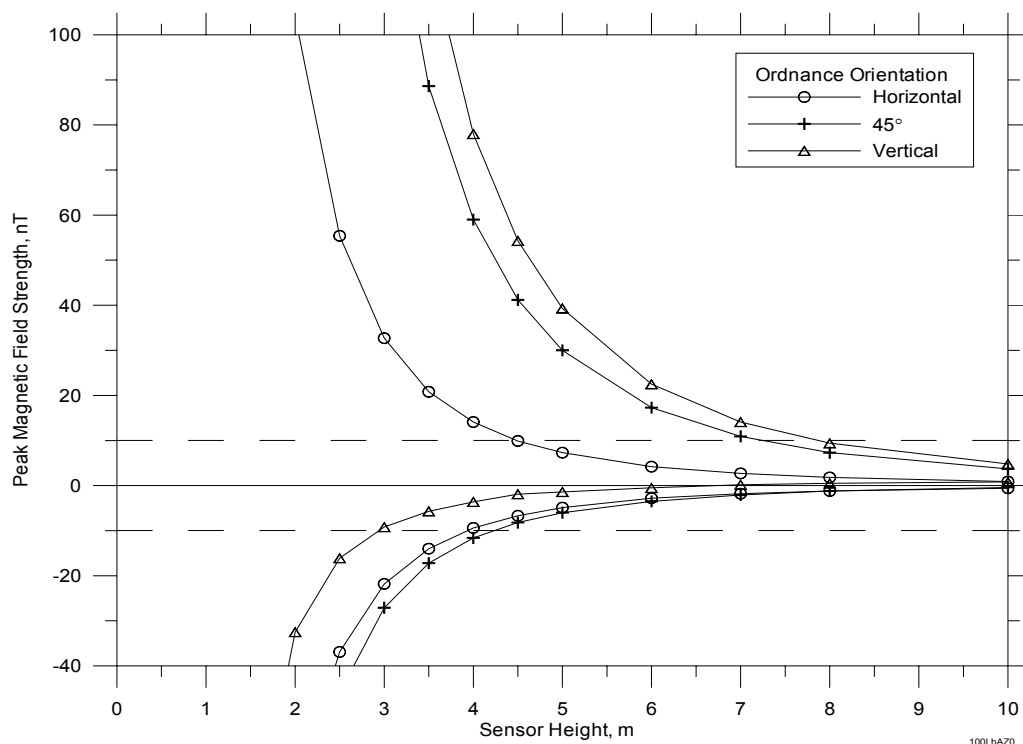
Fig. 23. Maximum and minimum total field anomalies for an M-38, 100-lb practice bomb, pointing toward magnetic north. Model calculations were made with MAGMOD. The dashed line represents a possible detection threshold, although a lower threshold will be appropriate for some sensor platforms and many sites with "quiet" magnetic background.

38

Measured responses from ordnance objects that are representative of a site, at a test grid, for example, may also be used to bracket the anticipated responses. These can be acquired with the ordnance in representative background geology, in order to characterize both the response and the background noise at once. It is useful to conduct modeling in advance of construction of such a test site in order to use optimal orientations for the UXO.

The range of depths and orientations of UXO at a site can be predicted or established using three approaches, one very conservative, one practical, and one "exact"; (1) "theoretical" maximum penetration depths can be determined using penetration equations, tables, or nomograms for vertical impact at maximum achievable velocity (e.g., Tarno and Butler 1986; Department of the Army 2000); (2) depths and orientations based on realistic penetration depths and recovery experience at a variety of locations; (3) penetration depth and path for specific cases using a sophisticated penetration code (e.g., Adley et al. 2003).

Penetration depths are dependent on soil and/or rock physical properties and heterogeneity, type of UXO (mass, length, diameter, nose shape, fins if present, etc.), and the details of the impact (velocity, angle of impact, angle of attack, spin, etc.). The type of soil/rock, vegetation, and soil moisture are some environmental factors that influence how deep an ordnance object will penetrate into the ground. Some general

Table 1. Selected portion of a tabulation of maximum penetration depths from Department of the Army (2000)

Ordnance Item	Depth of Penetration (ft)[1,2]		
	Sand	Loam	Clay
84 mm, M136 (AT4)	2.5	3.7	5.0
3.5" Rocket, M28	0.8	1.1	1.7
90 mm, M371A1	2.0	2.7	4.1
25 lb. Frag Bomb[3]	2.1	2.8	4.3
AN-M41A1 20 lb. Practice Bomb	5.0	6.6	10.0
105 mm, M1 (charge 7)	7.7	10.1	15.4
106 mm, M344A1	6.5	8.5	13.0
4.2" Mortar, M3 (max charge)	4.1	5.4	8.3
Dragon Guided Missile	0.9	1.1	1.7
155 mm, M107	14.0	16.4	28.0
8", M106 (charge 8)	16.4	24.2	36.9
M38A2 100 lb. Practice Bomb	8.6	11.3	15.2

[1]Penetration depths include the following "worst-case" conditions assumptions: impact velocity is equal to maximum velocity of round; impact is perpendicular to ground surface; munition decelerates subsurface in a straight line; munition does not deform upon impact. Typical penetration depth for any individual item will usually be significantly less.

[2]Actual detection depth may vary based on field conditions and be either lower or deeper.

[3]All bombs are assumed to have an impact velocity of 1135 feet per second.

[4]Maximum depth of penetration assuming a velocity of 500 fps.

observations have been noted for soils (Dept. of the Army 1986): (1) penetration depth decreases with increase in bulk density; (2) for materials having the same density, the finer the grain size the greater the penetration; (3) penetration depth increases with increasing water content. Geological factors such as frost heave, flooding, erosion and deposition, and human activities (agriculture, construction, recreation) can cause movement of ordnance object after its initial penetration.

An example of the conservative approach is shown in Table 1, where the maximum penetration depths for selected ordnance types (including the 100-lb practice bomb considered in Fig. 22) in three types of soil are tabulated. An example of the usefulness of the conservative, maximum penetration depths is shown in Fig. 24, where the total magnetic field magnitude versus depth for two orientations of a 105-mm projectile are compared to a nominal detection threshold and the maximum penetration depth for a typical soil. The maximum penetration depths correspond to case (d.) in Fig. 25, where the ordnance impacts vertically at maximum velocity.

The rest of the cases in Fig. 25 indicate an approach to practical penetration depth considerations that indicate a considerably reduced *vertical* depth of penetration for actual angles of impact and realistic impact velocities. Cases (a. – c.) indicate that ordnance will follow a curved trajectory ("J-hook" trajectory); and in the extreme case of grazing angles of impact (e.g., < 20°), the ordnance fragment will not remain buried at all, but will "breach or ricochet" and return to the surface. Two useful facts result from

Total Magnetic Field Calculations
Magnetic Field Inclination = 65°; Sensor Height = 0.5 m
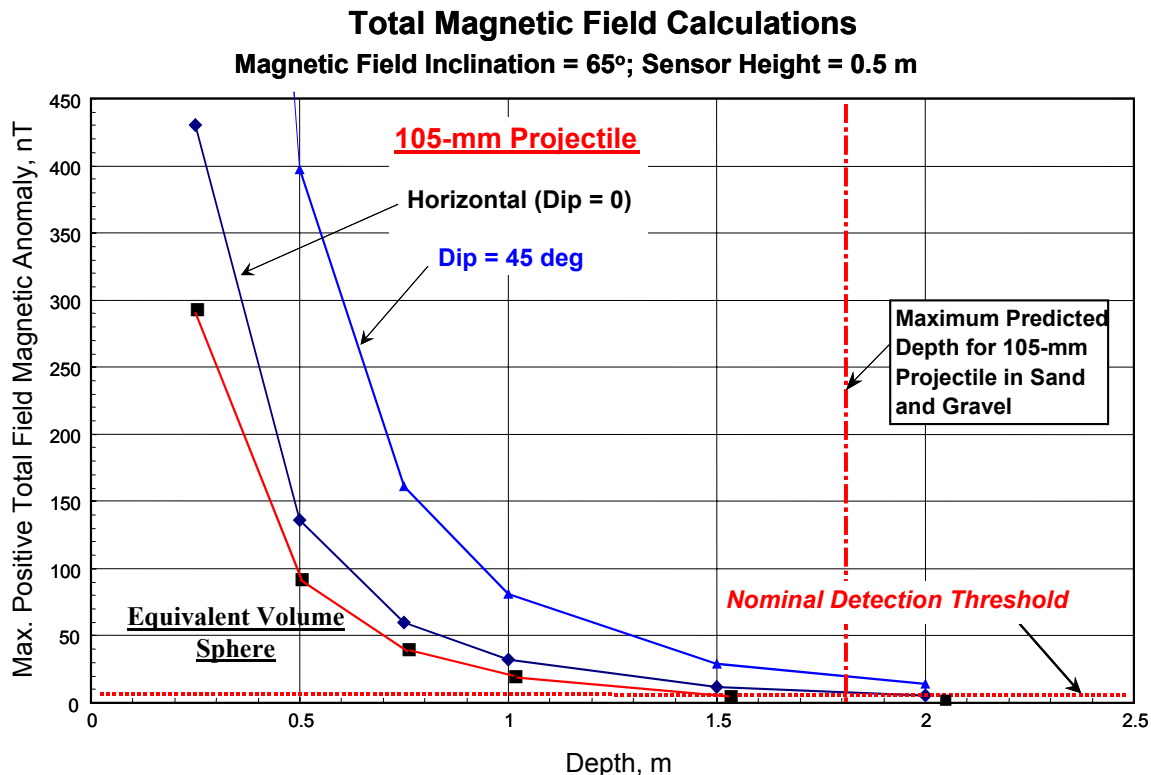
Fig. 24. Total magnetic field anomaly calculations for two orientations of a105-mm projectile, compared to a nominal detection threshold, the maximum predicted penetration depth for sand and gravel, and an equivalent sphere.
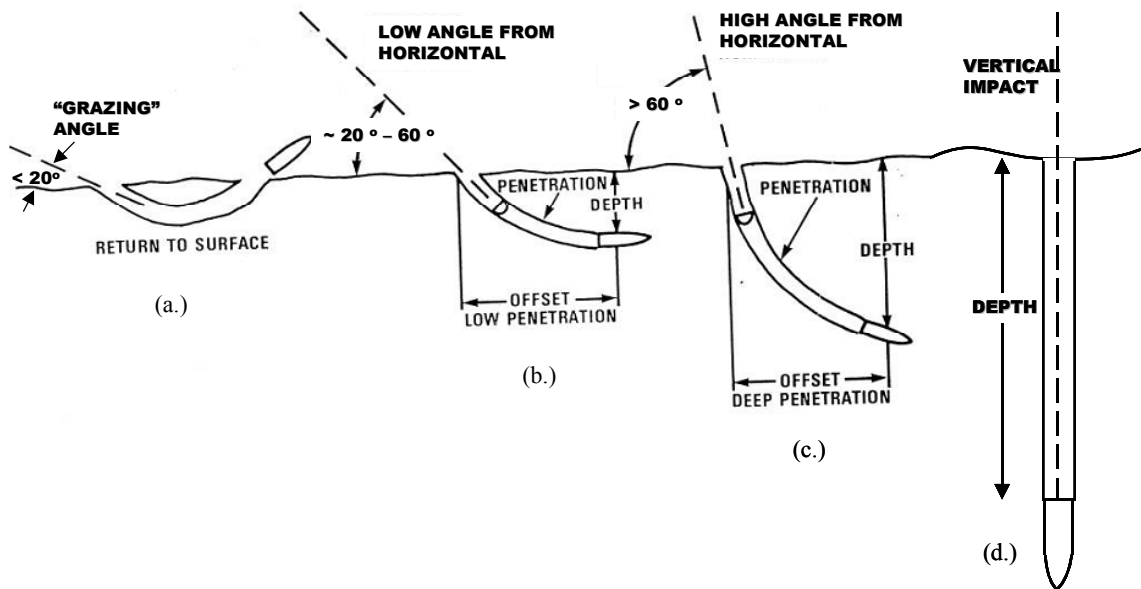
Fig. 25. Illustration of subsurface trajectories of UXO as a function of impact angle.
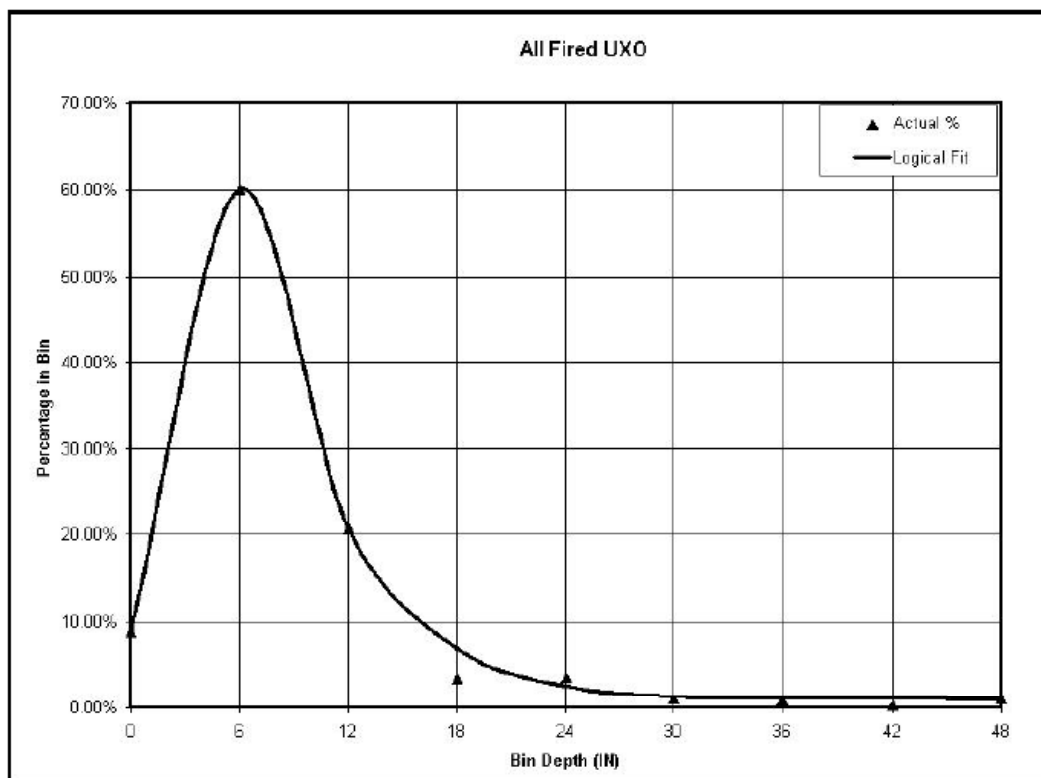
Note: The database used to develop this graph was populated predominantly with UXO items typically used by or in close support of ground troops. Large naval ordnance and large aerial bombs are under-represented.

Fig. 26. Plot of actual penetration depths based on extensive database of recovery depths of UXO at remediation sites (from Department of the Army 2000).

practical depth considerations: (1) actual penetration depths will always be less than the conservative, maximum depth; (2) the orientation of the ordnance item is observed intuitively and from recovery experience to be inclined predominantly < 45° relative to the surface and often is nearly horizontal. The second fact motivates the inclusion of curves for horizontal (0°) and 45° for modeling assessments such as Fig.22 and Table 1.

Practical penetration depth considerations are also indicated by UXO remediation recovery experience. For example, UXO cleanup at the former Fort Ord has resulted in an extensive database that indicates approximately 90% of ordnance pieces and ordnance scrap located within 0.3 m (12 inches) of the surface, and 98% located within 0.6 m (24 inches) of the surface   The Fort Ord experience is consistent with the data plotted in Figure 26 that is based on a much more extensive database from UXO remediation sites.

2.5 Geophysical to Point Pattern Data

Once a survey is designed, as described in Section 2.4, we assume that the survey will be conducted and geophysical data will be provided to guide subsequent excavation and analysis. We assume that some anomalies will be excavated to provide an improved basis for distinguishing between anomalies that are associated with UXO targets and those that are unrelated to UXO targets. In this section, we describe the tools and procedures for developing a point pattern data set from the geophysical data. This is not a trivial task, as it involves an understanding of the geophysical properties of UXO and the uncertainties associated with properly identifying the source of geophysical anomalies. If all geophysical anomalies above a selected threshold could be deemed to be UXO indicators, the problem would be trivial; unfortunately, this is never the case. Significant effort has been expended by ESTCP, SERDP, and other agencies toward improving this process at an individual UXO item scale, so as to reduce remediation costs.

The utility of geostatistical point pattern data for characterization of UXO sites as outlined in this report depends upon the interpretation of the sensor system response data, e.g., magnetic or electromagnetic data. The goal of this task was to develop statistical methods for classification of anomalies and to be able to distinguish between target-related objects (UXO, shrapnel) and target unrelated objects (geology, building materials, wire, scrap, etc.). Successful discrimination at this stage will clearly improve subsequent analyses that utilize anomalies as points in geostatistical analyses. Furthermore, feedback from ground truth data in the iteration between data collection and data analysis contributes to improved discrimination models.

Hard data on system performance is required in order to assign statistical properties to geophysical anomalies. For our example, the geophysical data from BBR that have been acquired with man-portable, MTADS ground-based, and ORAGS airborne magnetometer systems can be used to define system and platform performance. The geophysical anomalies mapped at these targets provide indications of the area dimensions of the targets. Selected excavations have been conducted at some of the targets. The statistical procedure would provide guidance on how to separate items that are target-related and target-unrelated.
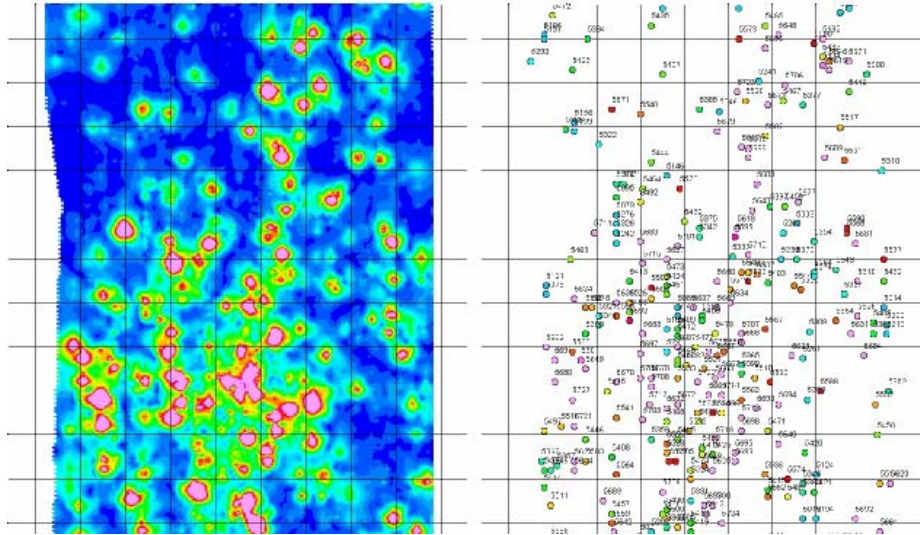
Fig. 27. Magnetic (analytic signal) map from a portion of Stronghold Table at BBR. The geophysical attributes of the anomalies can be used to distinguish target-related from target-unrelated anomalies. Here, the color-coding only indicates a categorization by anomaly amplitude, but we anticipate categorization based on other parameters, including +/- pole separation, wavelength, and spatial relationship to other anomalies. 20m grid cells.

Fig. 28. Airborne magnetic map (analytic signal) of Bombing Target 1 at BBR. Analysis of the distribution of anomalies and their spatial attributes will aid in setting survey design parameters. Results from digging such anomalies will establish spatial relationships between target-related and target-unrelated features (such as the fence line shown above). Area shown is approximately 600m by 600m.
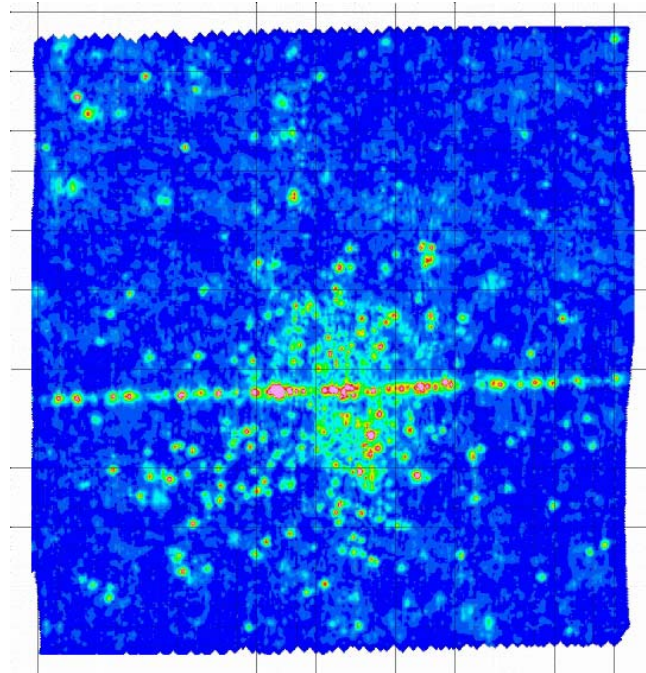
Although a mixture of intuition and statistical methods were used to select geophysical sampling locations at BBR, we can use the MTADS and airborne data that have been collected at BBR to represent the results from a first iteration of sampling. These include several airborne transects, MTADS surveys of two of the bombing targets in Fig. 3 and airborne surveys of each bombing target shown in Fig. 3.

Geophysical forward and inverse modeling as described in Section 2.4.3, can be used in interpretation of the geophysical data, in order to produce an estimated ordnance map from the geophysical map. Modeling can be used to bracket reasonable statistical parameters for anomalies, representing the probability that an anomaly is ordnance-related vs. non-ordnance-related. As noted previously in this document, it is important to distinguish this classification or discrimination system (ordnance vs. non-ordnance) from other discrimination efforts whose aim is to discern between different types of UXO, or to identify live UXO among inert UXO. An example of the application of a discrimination algorithm to the 60-mm example of Fig. 22b is shown in Fig. 29. Assigning probabilities to a ranked list of UXO-like targets is based on a goodness-of-fit measure or other objective criteria (e.g., Billings et al. 2002). This approach can be applied equally to magnetic or electromagnetic data with appropriate algorithms. The examples in this report are drawn from magnetic surveys simply due to the large volume of work and experience in this area.

2.5.1 Statistical Approach

The magnetometer-generated signature of a suspected ordnance item can be described as a two-dimensional "image" of the magnetic field surrounding the object. It has various characteristics of shape and amplitude that can be used to qualify or classify the source. Such inverse modeling (inversion) is an under defined (ill-posed) mathematical problem, making unique solutions impossible. Most discrimination efforts to date have focused on dipole moment modeling where an object's fitted dipole is compared to a library of ordnance dipoles. This approach requires the knowledge of the types of ordnance expected and their theoretical dipoles. The method was designed for small-scale ground surveys, and is arduous to implement in the case of airborne surveys where it may be necessary to examine thousands of anomalies in a timely manner. Furthermore, our experience thus far indicates automated versions of dipole model fitting programs have not yet proved reliable enough for use with large airborne data sets. The discrimination

Location	Recovered	Expected
Northing (m)	1.92	2.00
Easting (m)	2.04	2.00
Depth (m)	0.27	0.15

Orientation	Recovered	Expected
ϕ (degrees)	14.5	~ 0
θ (degrees)	51.7	~ 57

Transient Decay Parameters

k_1	10.89
α_1	0.014
β_1	0.94
γ_1	3.24
k_2	2.84
α_2	0.018
β_2	1.29
γ_2	2.96

Application of Discriminants

$\beta^{avr} = 1.11$ (>0.8) → *Ferrous Target*

$k_1/k_2 = 3.84$ (>1)
$\beta_1/\beta_2 = 0.73$ (<1) → *Rod-like Target*

⇒ *Algorithm correctly identifies shape and material property of target*

Fig. 29. Example of application of discrimination algorithm to parametric inversion results, see Figs. 21 and 22,for a 60-mm mortar; result indicates a ferrous, rod-like target, which must be listed as UXO-like in a target list. UXO probability ranking can be based on a goodness-of-fit criteria or similar objective consideration.

approaches we investigate here require preliminary calibration, as also do model fitting approaches, but can be employed in the field with minimal computing resources.

Magnetic detection systems have been employed at survey sites where the presence of ordnance objects is known or suspected. By design these systems respond to ferrous materials so that any anomaly that is defined by an elevated analytic signal could be caused by either an ordnance or non-ordnance ferrous object. Thus, on its own, the strength of the analytic signal is not an effective ordnance/non-ordnance discriminator. Note however, implicit in all discussions to follow, all anomalies subjected to classification have an analytic signal that exceeds a specified minimum threshold value.

Parameters, other than analytic signal can be derived from magnetic data, for example total magnetic field, magnetic peak value, magnetic low value, the separation between the magnetic peak and the analytic signal peak, magnetic peak-to-peak amplitude, width of analytic signal, width of total field peak, angle between magnetic north and the line connecting the maximum and minimum of the total field anomaly (theta), instrument height, estimated anomaly depth, and others. Although any single parameter may be only weakly correlated to the presence of UXO, when considered collectively they can be quite useful for anomaly characterization. Our approach to discrimination is an empirical approach that seeks traits measured by the suite of signal summaries that distinguish ordnance items (actually ordnance-related items, which includes ordnance fragments) from non-ordnance items. (Further work needs to be done to investigate, and possibly expand beyond a standard list, the types of signal summaries that can be used to this end.) Integral to this approach and to all supervised classification is training data, in our case a data set consisting of dig results and corresponding instrument signals. The objects found at a dig location make up a dig result. For purposes of discrimination the items are labeled as ordnance or non-ordnance; the non-ordnance group is further partitioned into subgroups that have distinguishable features, e.g. wire and scrap. We hope to enhance discrimination capability by providing categories of objects with relatively homogeneous characteristics even though our ultimate prediction will be either ordnance or non-ordnance. (This tactic could be taken for the ordnance group as well, although we do not do so because in the Badlands Bombing Range data set discussed here the ordnance group is fairly homogeneous; 90% of the items are M-38-related.)

There are various approaches to automated anomaly classification. In general, a discriminant rule relates to a division of the multi-dimensional feature space into disjoint regions that correspond to the set of predefined groups (Chapter 4, Gnanadesikan, 1977). An entity whose measured features fall into one of the regions is classified accordingly. For the approaches we discuss here the discriminant rule can be recast in terms of similarity or distance. For example, it would be reasonable to identify as ordnance any unknown whose measured parameters are similar to (close to, not far away from) those of ordnance. The approaches differ mainly according to how distance is calculated.

We can derive a measure of similarity to ordnance from the multiparameter means and covariances obtained from the training set consisting of the ordnance data alone. That measure is called the Mahalanobis distance. The concepts of "similarity to ordnance" and "distance from ordnance" can be visualized most simply for the two-parameter case. Consider a plot of hypothetical data, analytic signal versus magnetic low value, in Figure . If the two parameters were uncorrelated and had equal variances, the cloud of points on the plot would appear to vary about the mean point (shown as a filled circle)

approximately equally in all directions. Then a point on the plot that is "similar to ordnance" or "close to ordnance" would be close to the mean point in the usual Euclidean distance sense. This is clearly not the case in the (equal x, y scales) plot of Fig. 30a, where there is pattern of negative correlation between the two parameters and variability in the analytic signal dimension greater than in the magnetic low value dimension. Euclidean distance between points on this plot does not fit with our sense of similarity in this case. For example, the point P1 at (20, -8) on this plot is situated more like other ordnance points than point P2 at (12, 8) even though the Euclidean distance of either point from the mean is approximately the same. The calculation of Mahalanobis distance takes into account, the variances and covariances. Mahalanobis distance between points in the original scale, such as in Fig. 30a, is equivalent to Euclidean distance between points such as P1 and P2 in a transformed representation, such as in Fig. 30b, where the new dimensions are uncorrelated and have equal variance. The notions of Mahalanobis distance and transformation to uncorrelated variables extend to dimensions greater than two.
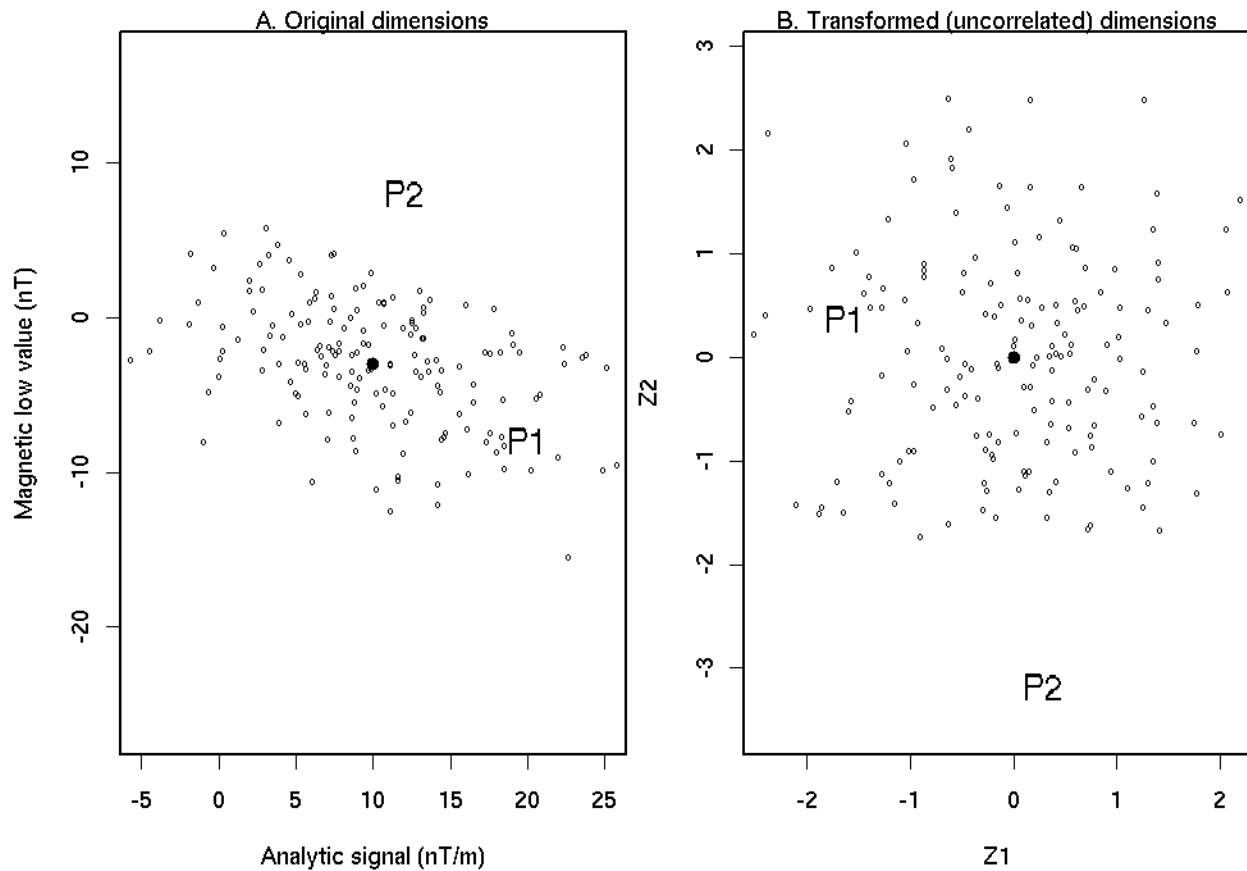
Fig. 30. Points and Euclidean distance in original and transformed dimensions.

We can calculate the Mahalanobis distance of any anomaly from the ordnance mean and order the distances from smallest to largest as a way to prioritize anomalies for follow up. We can choose a cutoff distance for discrimination from the receiver operator curve (ROC) derived from the training data set, assessing the trade-offs between increasing true positive rate (TPR) and increasing false positive rate (FPR). Available resources, time and money, will weigh heavily in the decision. We can in general calculate the Mahalanobis distance of any anomaly from any other group mean, e.g., scrap mean, in a similar fashion. We could even use this measure when training data were available only on ordnance items, but rather than deriving a cutoff from a ROC curve we would perhaps determine a upper tolerance bound for the ordnance population as a cutoff. See Pepe (2000) for description of ROC methodology.

A ROC curve is a useful diagnostic display to assess two-group classifiers. Consider classifying anomalies as ordnance when the analytic signal exceeds some cutoff. The ROC curve shown in Fig. 31 pertains to data collected at the Badlands Bombing Range (discussed below). Each point on the curve is linked to an analytic signal cutoff value. For example, the point (0.812, 0.95) corresponds to classifying an anomaly as ordnance if the analytic signal exceeds 4.2nT/m. In this data set 95% of the ordnance and 81.2% of the non-ordnance exceed this value. The point (0.5, 0.594) corresponds to a 7.5nT/m cutoff.
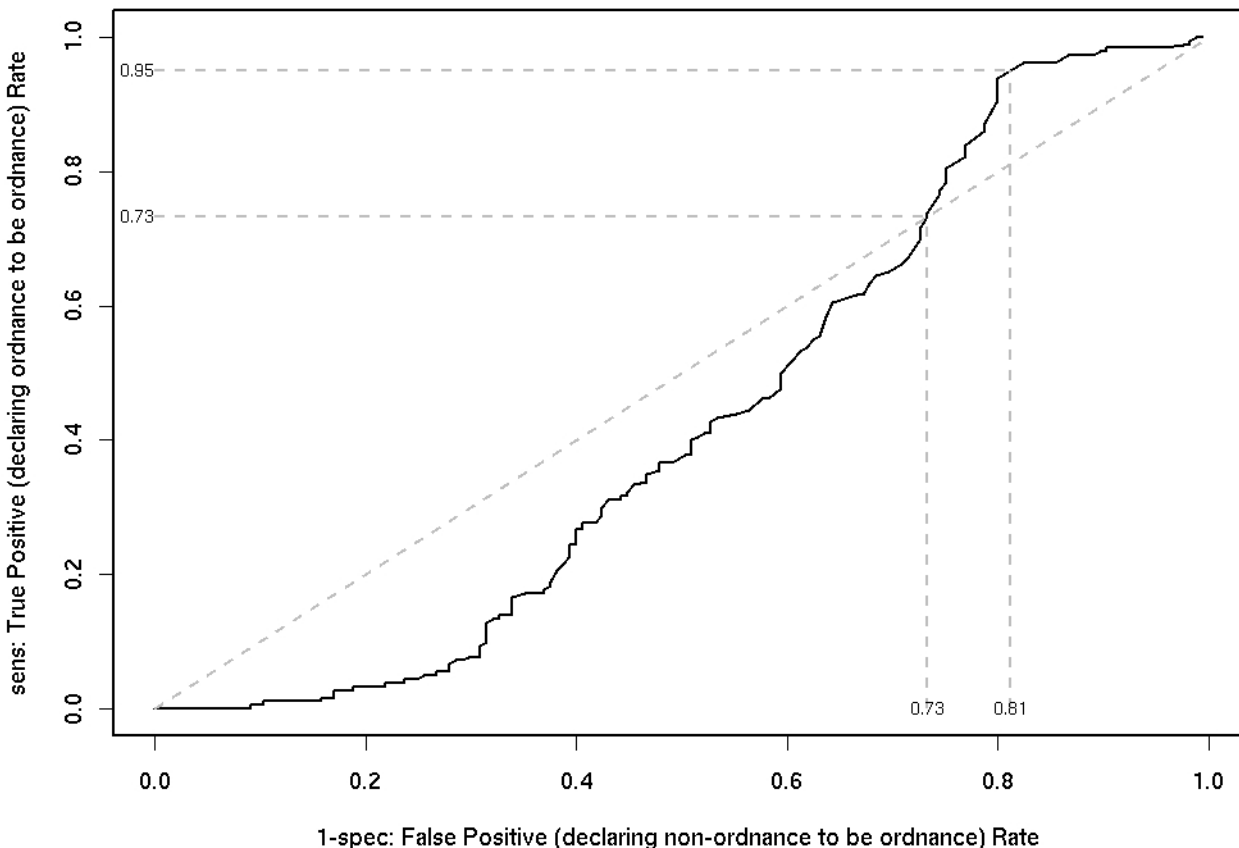
Fig. 31. ROC curve showing the rate of true positives versus false positives for a list of anomalies with ordering based on analytic signal (x ≤ cutoff).

The curve in Fig. 31 shares a property with all ROC curves in that it passes from (0, 0) to (1, 1). For this example, the (0, 0) point would imply that we classify all anomalies as non-ordnance and therefore get all of the non-ordnance correct but we get none of the ordnance correct. The (1, 1) point is where everything is classified as ordnance and therefore we get all ordnance correct but get all non-ordnance wrong. The diagonal line connecting (0, 0) and (1, 1) corresponds to a random classifier, i.e., a classifier that randomly guesses ordnance, non-ordnance by flipping a fair coin. Note that the ROC curve is below this diagonal line for the TPR up to approximately 0.73 indicating that for some cutoff values the classification rule does worse than classifying ordnance and non-ordnance by flipping a coin. Note however, that we could modify this classification rule to be better than a random rule overall. For example, setting an upper cutoff threshold that eliminates the first 30% of the false positives results in a classification rule that does much better than a random rule (since most of the very large responses are too large to be ordnance). It would, however, result in approximately 8% false negative responses – a parameter not captured in the ROC curve. Other rules, such as the multi-parametric system described above or the multivariate system below would have completely different ROC curves.

A perfect classifier would be a step function rising from (0, 0) to (0, 1) and flat to (1, 1). Thus, in general, a classifier whose ROC curve rises sharply and achieves its maximum quickly is to be preferred. A traditional method for comparing classifiers is the area under the ROC curve (AUC). The perfect classifier has AUC 1 so that classifiers can be easily compared using this measure. It is likely in practice that a classifier would not be operated at FPR's that are very large. In this case, by specifying an upper bound on the FPR, we could evaluate classifiers over the limited FPR range by comparing the partial area under the ROC curve (pAUC). Continuing the example, the AUC for the analytic signal classifier is 0.43 and if we specified an upper FPR of 0.8 the pAUC is 0.24.

Linear discriminant analysis (LDA) is another approach to classification where a discrimination rule is derived from a training data set made up from the exhaustive list of groups/categories of interest. The rule classifies a new item according to its (multivariate) similarity to a group, i.e., LDA classifies to the predefined group the item is closest to. You can think of LDA classifications based on Euclidean distances after a mathematical transformation, as described above for Mahalanobis distance. The transformation used in LDA is different from the one used for Mahalanobis distance. Other discrimination approaches can be derived as variants of LDA or Mahalanobis distance, e.g., calculate Mahalanobis distance from each group separately and classify into the closest group. One disadvantage of traditional LDA and this variant of Mahalanobis distance from each group is that all groups must be predefined and have training data. See Section 4.2.1 of Gnanadeskikan (1977) for a general discussion of distance measures as related to classification.

In this report we focus on two classification procedures, Mahalanobis distance from ordnance and a variant of the two-group traditional LDA denoted here by LDAv. In LDAv we perform a usual two-group LDA using a training set with items known as either ordnance or non-ordnance (i.e., the further refinement of non-ordnance into subgroups is ignored). In the two-group case LDA simplifies to classifying into one group or the other based upon a single transformed variable, x which is a special weighted sum of the original variables. The special weights are determined so that the t-statistic testing group difference is maximized using data on the transformed x (Panel on Discriminant

Analysis, Classification, and Clustering 1989). A cutoff value in the transformed scale, called the first discriminant coordinate (LD1), determines whether an item is classified into one group or the other. The method, which the cutoff value is chosen, is how LDAv differs from usual LDA. We use the ROC curve derived from the training set to assess and choose LD1 cutoff values.

Implicit in the discussion of multivariate classification is that there is a defined set of variables to be used. We identified a number of parameters that could be calculated from the magnetics data without regard to their utility for classification. We now want to select effective discriminators. If we include too many variables we may do very well when classifying the training data set but do worse when classifying new data. Therefore, we would like to strike a balance between the number of variables and good discrimination. We performed a two-stage screening of the variables. In the first stage we identified and dropped variables that were essentially redundant. In the second stage we compiled the final collection of variables from lists obtained using regression analysis techniques.

We couched our search for the best list of variables within the framework of classical LDA. Our primary interest was to discriminate ordnance from non-ordnance where we had several types of non-ordnance items. We separated the non-ordnance items into groups to take advantage of signal characteristics that may exist and may improve discrimination. We applied standard variable selection techniques to the two-group discrimination problems, ordnance versus each of dirt, scrap and wire, and then determined the final set of variables from the combined list. Parameters appearing in multiple lists and consistently within discrimination sets of various sizes were favored for the final list. The two-group LDA can be recast into a regression setting (Panel on Discriminant Analysis, Classification, and Clustering 1989) where group membership, recoded as a binary variable (say 0 for one group and 1 for the other group), is the dependent variable and the potential discriminators are the independent variables. Then an all-subsets regression routine can help assess the value of classifier sets.

2.5.2 Analysis of Badlands Bombing Range Data

Data were acquired at the Badlands Bombing Range (BBR) in South Dakota, using a helicopter boom-mounted magnetic detection system during 2000 and 2001. The site, instrumentation, and ground follow-up is described ORNL, 2000. The raw data were gridded to a cell size of 1m and post-processed with Geosoft Oasis Montaj software (Geosoft, 2003), to compensate for instrumentation system and configuration influences and other systematic effects, to calculate the analytic signal, and to select anomalies analyzed here. Anomalies were chosen based upon a threshold intensity of the analytic signal peak (analytic signal > 1.2 nT/m to 1.5 nT/m). Further statistical analyses were performed using the open-source R statistics package (Ihaka and Gentleman, 1996). Some, but not all of the surveyed sites, had dig information that could be associated with anomalies. Dig results and magnetic data were paired according to physical location, easting and northing coordinates. The closest dig result within 3 meters of a magnetic anomaly was considered a match.

A suite of summaries for the training set was calculated for each anomaly. Each variable considered in this investigation is listed in Table 2, along with a brief description.

Table 2. List and Description of Variables Used in the Statistical Analyses

Variable	Description (units)
*as.approx	Analytic signal approximation(nT/m)
*as.grid	Analytic signal(nT/m)
*depth	Calculated anomaly depth(m)
inst.height	Instrument Height(m)
mag.pp	Magnetic peak-to-peak amplitude(nT)
*mag.pp.sep	Magnetic peak separation(m)
*maglow.value	Magnetic low value(nT)
magpeak.as.sep	Separation of magnetic peak and analytic signal(m)
magpeak.value	Magnetic peak value(nT)
*theta	Offset angle of total field anomaly from magnetic north(deg)
*width.as	Width of analytic signal(m)
*width.tf.peak	Width of total field peak(m)
ihpd	Instrument height + calculated anomaly depth(m)
x.as.peak	Easting of as.grid(m)
x.maglow	Easting of maglow.value(m)
x.magpeak	Easting of magpeak.value(m)
y.as.peak	Northing of as.grid(m)
y.maglow	Northing of maglow.value(m)
y.magpeak	Northing of magpeak.value(m)
*rat.asg.asa	Ratio, as.grid:as.approx
rat.mpass.ihpd	Ratio, magpeak.as.sep:ihpd
*rat.mpass.mpps	Ratio, magpeak.as.sep:mag.pp.sep
*rat.mpass.was	Ratio, magpeak.as.sep:width.as
*rat.mpass.wtfp	Ratio, magpeak.as.sep:width.tf.peak
rat.mpps.ihpd	Ratio, mag.pp.sep:ihpd
*rat.mpps.was	Ratio, mag.pp.sep:width.as
*rat.mpps.wtfp	Ratio, mag.pp.sep:width.tf
*rat.mpv.mpp	Ratio, magpeak.value:mag.pp
rat.was.ihpd	Ratio, width.as:ihpd
rat.wtfp.ihpd	Ratio, width.tf.peak:ihpd
*rat.wtfp.was	Ratio, width.tf.peak:width.as

*Indicates variables considered as potential candidates for discrimination between ordnance items and non-ordnance items.

The variables are either summaries calculated from gridded raw data using Geosoft Oasis Montaj (Geosoft, 2003) software or derived variables, functions, mostly ratios, of those summaries. The signal characteristics of an item vary with the item's distance from the sensor. We attempt to remove the effect of item-to-sensor distance on an anomaly's signal through these derived variables and gain additional insight on discriminating ordnance items from non-ordnance items. (Further research into variable selection is part of a new study proposal submitted by ORNL.)

Table 3 contains the dig list grouped into broad categories. A large number of the

Table 3. Badlands Bombing Range Anomaly Descriptions

Anomaly group Anomaly description	Frequency
==	
Soil	
Magnetized soil	25
Magnetized rock	2
Ordnance	
250 lb.sand filled practice bomb, AN-M57	1
4lb incendiary pieces-fuse(burnt) & frag	2
Bomb Frag/Bomb fin	8
Bomb Frag & Barbed Wire	2
M50 Series Inc. bomblet, Blown in Place	1
M38 pieces/fin/frag	8
M38/M38 (Backhoe)/M38 & & Functioned Bomb Fuze, M100 Series	144
M38 100lb Practice Bomb	8
M38 with Live Spotting Charge/M38 link of (3) .50 cal (rds)	2
Strongback from Incendiary Dispenser	4
Scrap	
Bed Spring	1
Can/Can lid/Cooler top/Paint can/20mm ammo can	8
Car parts/Buried car/Old farm implement	5
Corrugated tin/Wash tub/Barrel bottom/Bucket	5
Culvert/Culvert under paved road	2
Hay rake/Plow blade/Rake tooth	5
Metal/Rusty metal pieces	2
Metal Wagon Wheel/Steel loop	2
Radar Reflector Target	2
Rebar/Steel bar/Metal stake/Metal rod	12
Steel posts/Fence post/Fence and post	10
Steel pipe/Exhaust pipe	2
Target anchor/Chain/	7
Wire	
Wire	18
Barbed wire/woven wire	51
Barbed wire and metal/Pipe and wire	2
Old fence post & posts w/padlock	1
Guide wire post anchor/Telephone pole and cable	3
Unknown or nothing detected	4
No dig results	1178
==	
	1527 (tot)

anomalies have no associated dig results because no follow-up digs have been done. We focused on the four main groupings of dug anomalies, ordnance/ordnance-related (ord, n=180, 162 are M-38's or M-38 fragments), magnetic rock or dirt struck by lightning (dirt, n=27), scrap metal (scrap, n=63), and wire (wire, n=75) as the training set (n=345) for developing discrimination tools. The statistical analysis of that data set is described here.

Overall and group statistical summaries are presented in Appendix 2 for each of the variables listed in Table 2. The statistics calculated are n total (n), the number equal to 0 (n=0), mean (Mean), standard deviation (SD), minimum (Min), order p quantiles for p=0.5, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95 (0.5, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, respectively), and maximum (Max). Boxplots in Fig. 32 provide visual comparisons of univariate distributions with regard to anomaly group. The lower and upper extents of a box correspond to the interquartile range, the 0.25 and 0.75 quantiles, the line within the box corresponds to the 0.5 quantile (median), and the dashed lines extend to the most extreme data points which are no more than 1.5 times the interquartile range from the box. Data values more extreme are plotted individually. Box width is proportional to the square root of the number of observations in the group. To facilitate comparisons of the bulk of the data, some of the plots are clipped either above or below indicated in a plot's marginal text by the cutoff value(s) used and the resulting number of data points from each group not shown.

Some of the individual summary parameters show potential for discriminating between ordnance items and non-ordnance items. Consider, for example, magnetic low value, analytic signal, the ratio of anomaly magnetic peak value to magnetic peak-to-peak, and separation between magnetic peak and analytic signal peak, respectively. These plots show and the summaries in Appendix 2 confirm, for example, that 95% of the ordnance items of the types encountered here exhibit analytic signals below 21 nT/m, will tend to have small (in absolute value) magnetic low values (95% are greater than -11 nT/M), and tend to have shorter distances between the magnetic peak and analytic signal than non-ordnance (95% are less than 3.2 m). Furthermore the ratio of positive magnetic peak to magnetic peak-to-peak is a fair discriminator of ordnance items from dirt, the bulk of ordnance values are larger than the bulk of the dirt values.
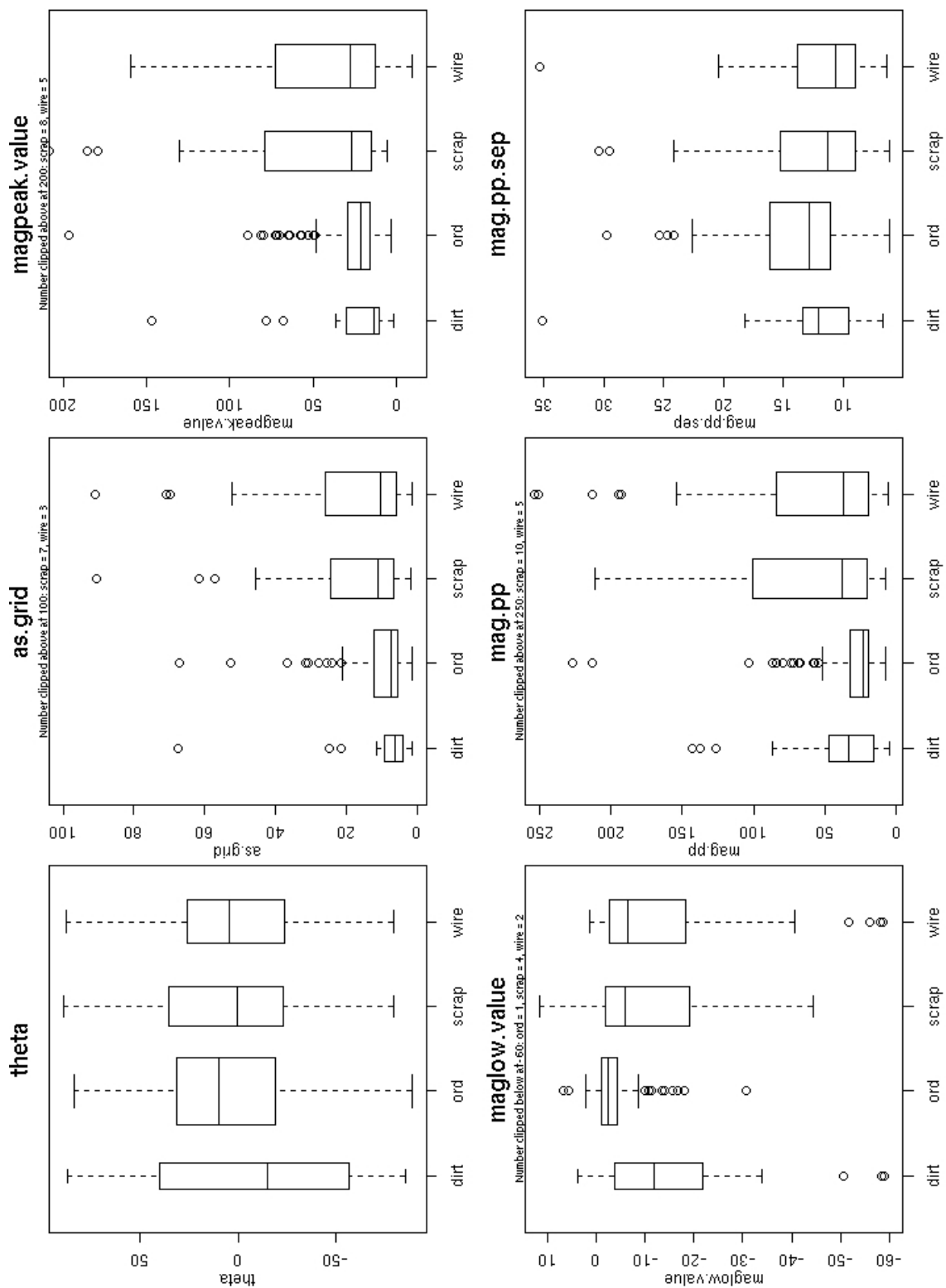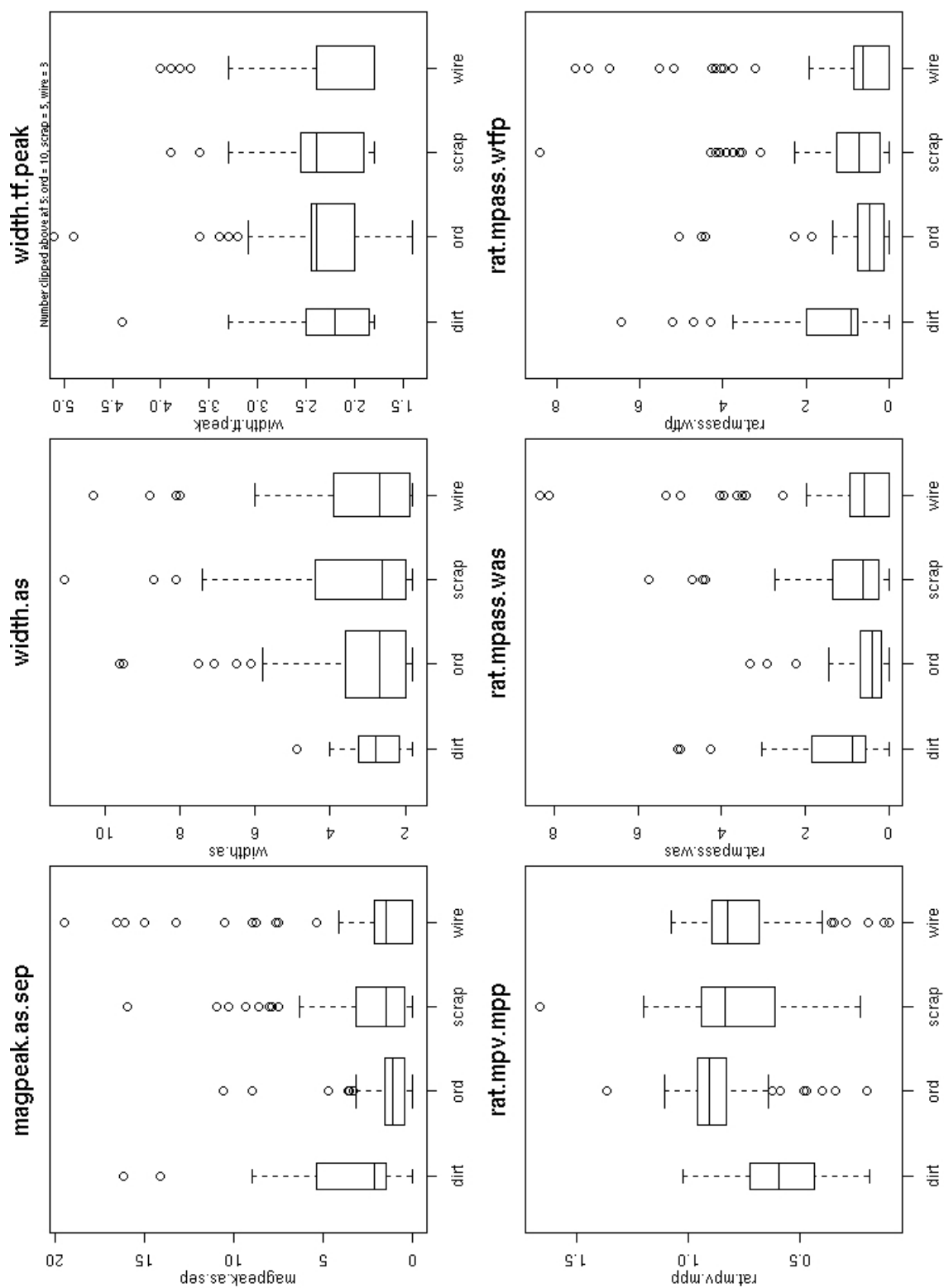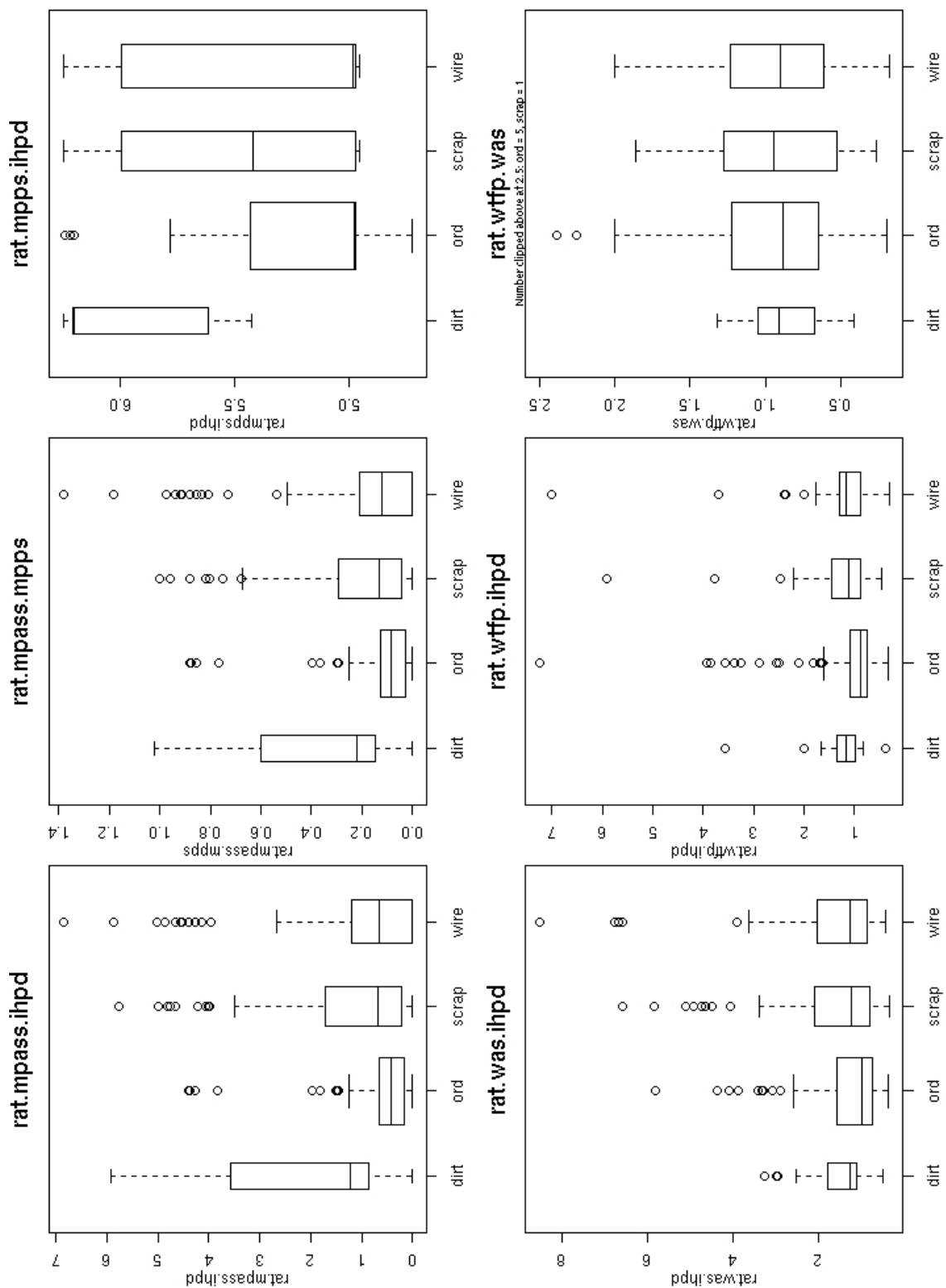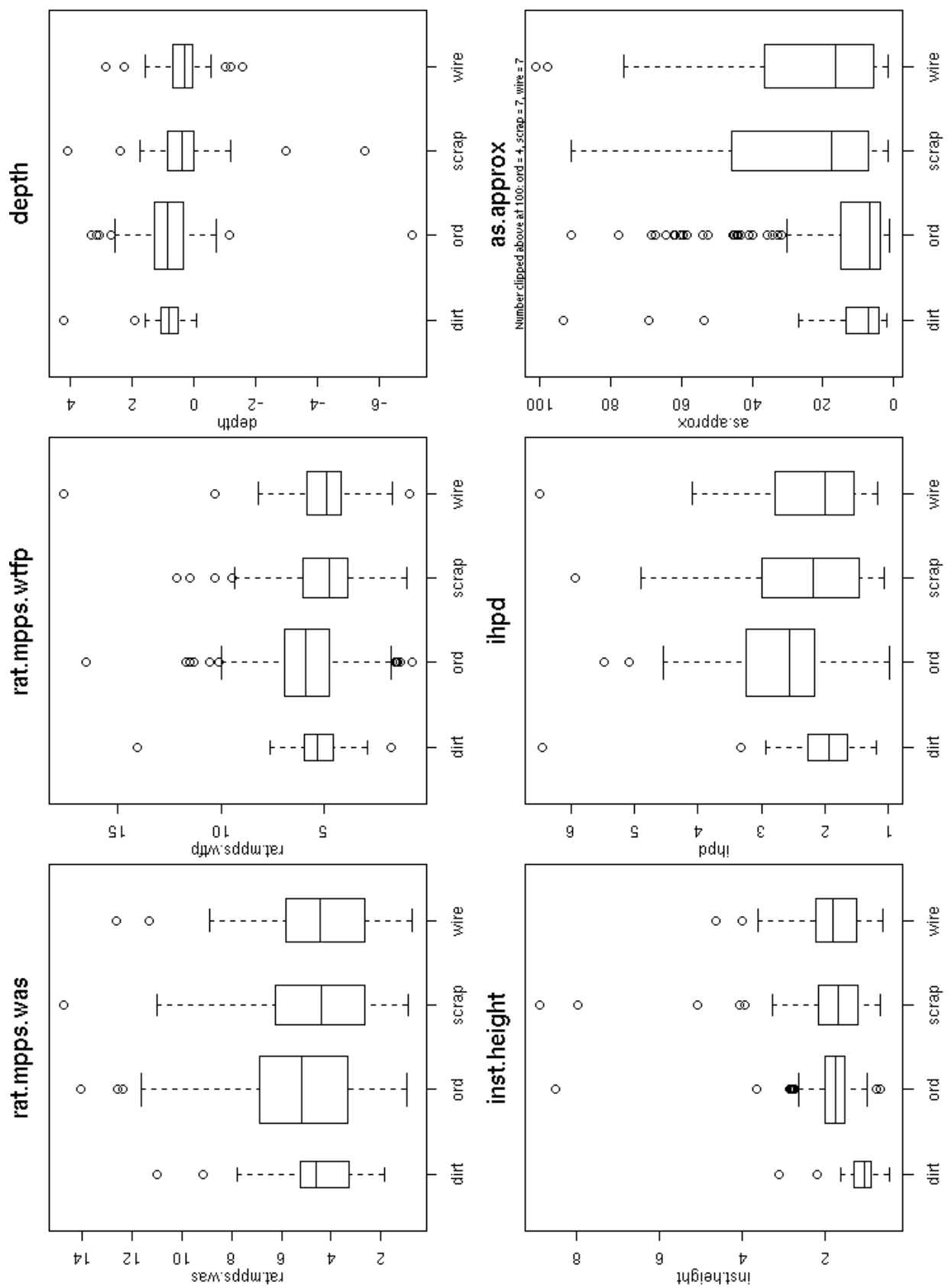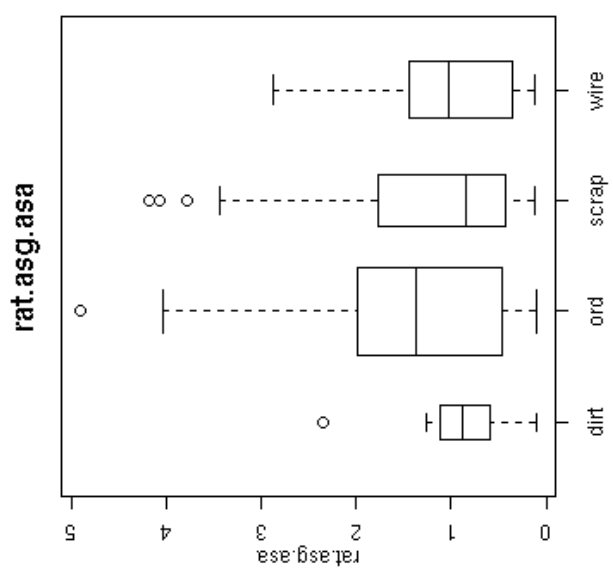
Fig. 32. Boxplots comparing anomaly groups with respect to variables.

Collectively, a suite of parameters should improve our ability to discriminate between ordnance items and non-ordnance items. In each of the approaches to multi-parameter discrimination we must specify that set of parameters. Some of the variables listed in Table 2, provide redundant information and some may not have utility for discrimination at all. If we include too many variables we may do very well when classifying the training data set but do worse when classifying new data. We would like to strike a balance between the number of variables and good discrimination.

As a first level of screening for variables to be used for discrimination, we viewed various x,y plots and assessed correlations among variables (not presented here). We identified a number of redundant parameters. When selecting among alternatives we preferred variables that have some theoretical basis for discrimination (e.g., Nelson, et al. 1998) suggest that the offset angle for UXO may be within say 35-50 degrees of the earth's field because of shock demagnetization), variables that may account directly for instrument height and item depth (e.g., ratios), and variables that have desirable, interpretable properties (e.g., analytic signal). The analytic signal serves as a surrogate for both the magnetic peak value and magnetic peak-to-peak amplitude, however the ratio of magnetic peak value to magnetic peak-to-peak amplitude was retained. Instrument height plus calculated item depth, was found to be highly correlated with the magnetic peak-to-peak separation; the latter was retained as a variable and in ratios, rather than the former. The magnetic peak, analytic signal separation variable was highly correlated with all ratios where it was the numerator; one ratio, the ratio of magnetic peak, analytic signal separation to width of the analytic signal was retained. The list of potential variables for discrimination after the initial screening, are indicated in Table 2.

We sought the best list of variables for discrimination of ordnance from non-ordnance using two-group LDA, ordnance versus each of dirt, scrap, and wire. We cast LDA as a regression problem and used variable selection techniques to arrive at candidate sets. The final set was derived from the best sets from each two-group LDA. The top single-variable discriminators for ordnance and the individual non-ordnance categories were fairly consistent, either analytic signal (as.grid), magnetic low value (maglow.value), ratio of magnetic peak value to magnetic peak-to-peak (rat.mpv.mpp), or ratio of magnetic peak, analytic signal separation to width of analytic signal (rat.mpass.was). We settled upon a six-variable discrimination set after additional subsets were evaluated, all including the four variables just listed. The two additional variables added were offset angle (theta) and width of analytic signal (width.as). Although this six-variable list is not the overall best list for each two-group discrimination problem, it is among the best for each.

First, consider Mahalanobis distance of anomalies from the ordnance mean for the different groups of anomalies. Figure 33 shows boxplots comparing distance from the ordnance mean for the four groups. Clearly there is a benefit to considering parameters collectively; it appears that approximately 75% of the calculated distances for the ordnance group are below 2.4 units whereas approximately 75% of the distances in the other groups are above that level. Figure 34 shows a ROC curve for Mahalanobis distance. This classifier classifies anomalies according to a cutoff rule; larger distances are more consistent with non-ordnance. One possible rule identified on the plot corresponds to a cutoff distance of 4.41 units resulting in 95% of the ordnance correctly identified and 63.6% of the non-ordnance incorrectly identified as ordnance. The pAUC for a 0.8 FPR is 0.61 for this classifier, much improved over the analytic signal classifier discussed earlier.
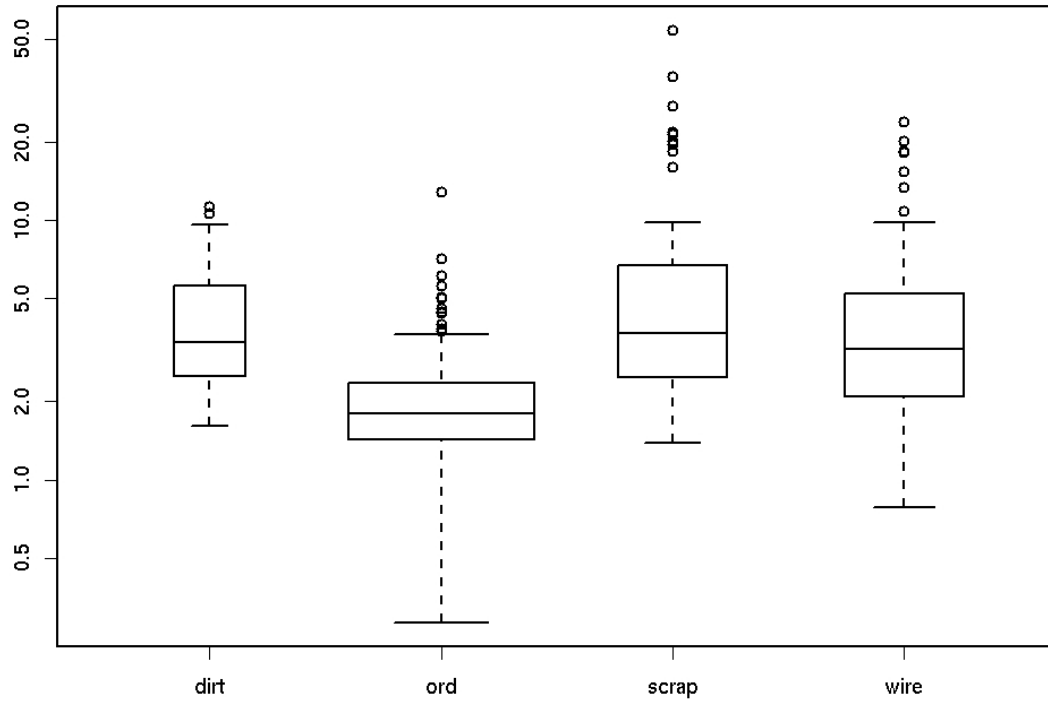
Fig. 33. Boxplots comparing anomaly group with respect to Mahalanobis distance.
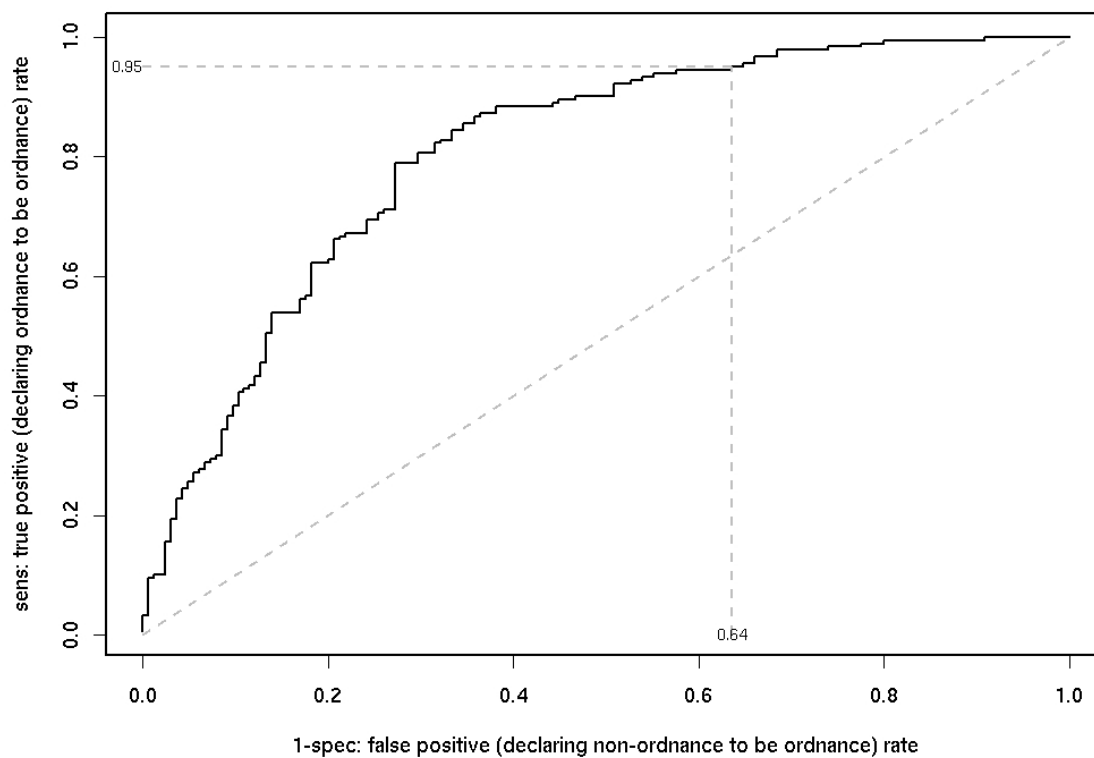
Fig. 34. ROC curve showing the rate of true positives versus false positives for a list of anomalies with ordering based on Mahalanobis distance (x < cutoff).

Now, consider anomalies with respect to LD1, the first linear discriminant coordinate in the LDAv classifier. Figure 35 shows boxplots comparing LD1 in the four groups. Again, as was observed with the Mahalanobis distance, separation between ordnance and non-ordnance groups in this transformed variable is enhanced when compared to that for individual parameters as seen in Fig. 32. Figure 36 shows the LDAv classifier ROC curve. This classifier classifies anomalies as ordnance when LD1 exceeds a cutoff value, i.e., larger LD1 values are more likely with ordnance items than non-ordnance items. The rule identified on the plot classifies anomalies as ordnance when LD1 exceeds -0.28 units, which results in 95% of the ordnance correctly identified and 50.3% of the non-ordnance incorrectly identified as ordnance. The pAUC for a 0.8 FPR is 0.59 for the LDAv classifier.

2.5.3 Summary

Parameters, other than analytic signal, have utility for discriminating between ordnance items and non-ordnance items; however, the multivariate nature of magnetic anomalies can be exploited for improved discrimination. There are a variety of classification techniques; all require a training set for calibration, representative of the area, ordnance types, and measurement technology to be applied.

We presented two classifiers that utilize multivariate information found in the magnetic signal; both can be tuned using ROC curves. One classifier is based upon Mahalanobis distance, an intuitive measure of similarity. Ordnance training data are required in order to develop a framework for measuring similarity to ordnance, but training data are not required for non-ordnance items unless we want to tune the classifier using ROC curves. LDAv is the other classifier, which is a variant of traditional linear discriminant analysis. It classifies according to a transformed variable LD1, which is the first linear discriminant coordinate. It too is quite intuitive in that the transformation maximizes the separation between ordnance and non-ordnance items. A full training set of ordnance and non-ordnance items is required.

A set of discriminator variables was selected in two stages from the full suite of signal parameters. In the first stage we dropped redundant variables, retaining those that are easily interpretable or theoretically meaningful. We recast two-group LDA, ordnance items versus each of dirt, scrap and wire, as a regression problem and used variables selection techniques to arrive at the final set of discriminators. For the BBR data we selected a six-parameter set, analytic signal (as.grid), magnetic low value (maglow.value), ratio of magnetic peak value to magnetic peak-to-peak (rat.mpv.mpp),
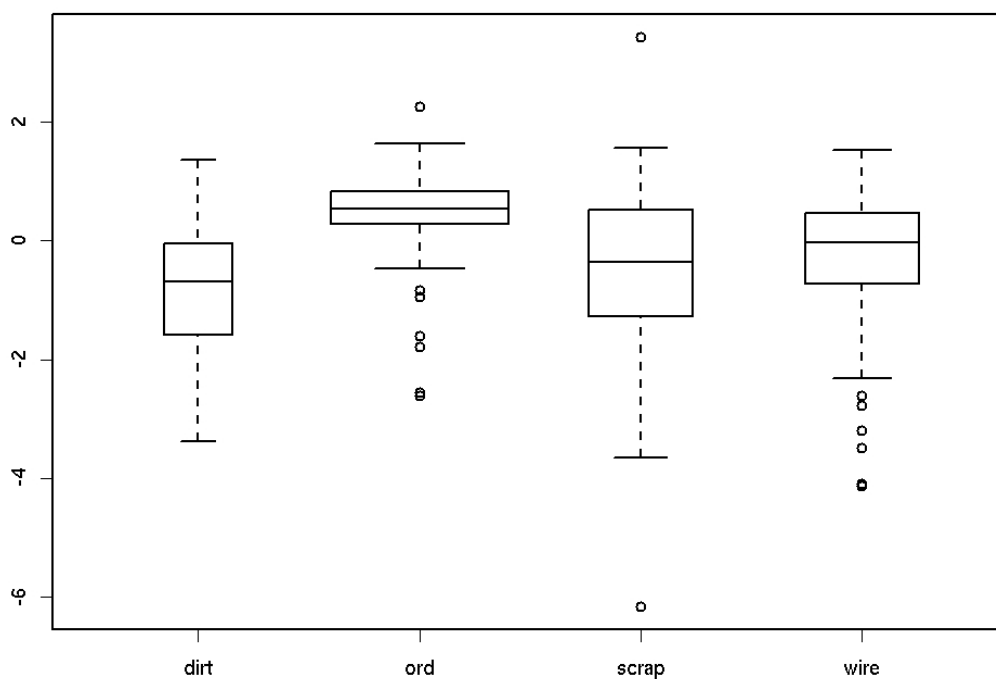
Fig. 35. Boxplots comparing anomaly groups with respect to the first discriminant coordinate.
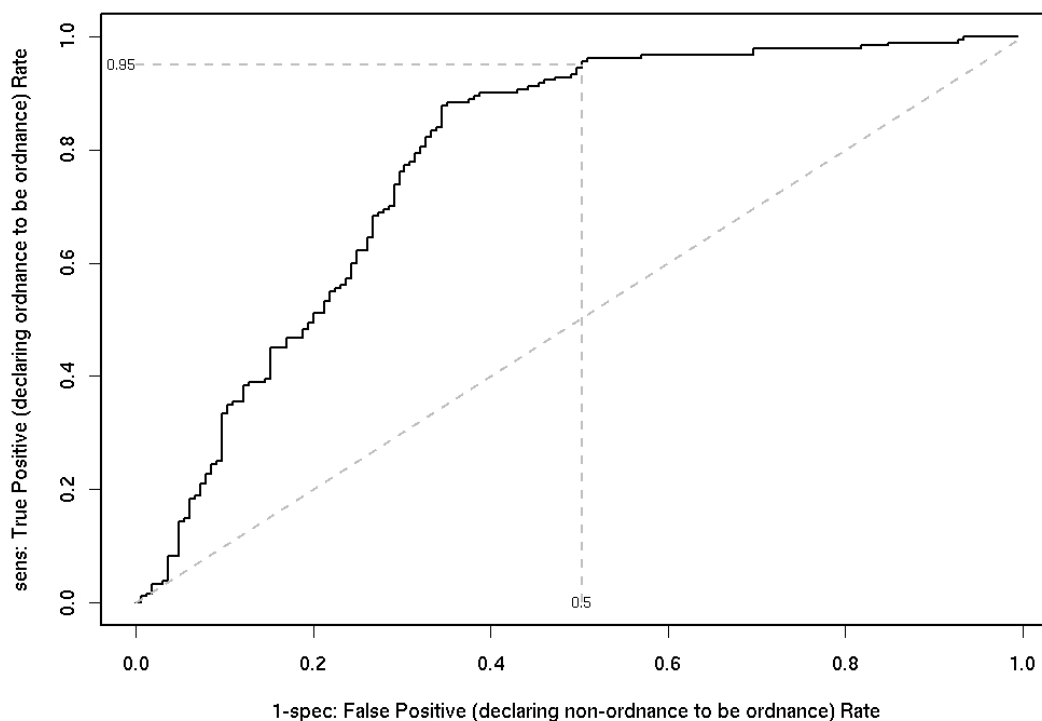
Fig. 36. ROC curve showing the rate of true positives versus false positives for a list of anomalies with ordering based on first discriminant coordinate.

ratio of magnetic peak, analytical separation to width of analytic signal (rat.mpass.was), offset angle (theta) and width of analytic signal (width.as).

In the BBR data analysis, the Mahalanobis distance and LDAv classifiers had similar performance, but in general this may not be the case. One reason for a difference is that we do not take into account direction. When Mahalanobis distance is calculated, an ordnance item quite distant from the ordnance mean could be even further distant from the non-ordnance groups, but be classified as non-ordnance. We would expect the LDAv classifier to be better in situations when covariance structures are the same in the ordnance and non-ordnance groups, because by design, LDA finds the best separation between the populations. Performance will be degraded, however, when covariances are heterogeneous, but LDA is known to be quite robust to assumptions. Both classifiers discussed here are definite improvements over classifiers using single parameters such as analytic signal. Finally, once classification of detection instrument signals is complete, the locations of those predicted to be ordnance-related is utilized in the next analysis stage.

2.6 Construction of OIM from Point Pattern Data

Although the point pattern locations themselves can be used in nearest neighbor distance estimation methods, when estimating the intensity of an inhomogeneous Poisson process from an incomplete sample, the mathematics becomes intractable (Diggle 2003). However, methods exist for Poisson count data. For this reason, the point pattern of the ordnance location data is converted to counts on a grid, as shown in Figs. 4 and 5, and is used in the estimation methods described in this section.

2.6.1 Estimating the OIM

Estimation procedures for point pattern data are discussed in Cressie (1991), Stoyan et al. (1995), and Diggle (2003). The Neyman-Scott process, introduced in Section 1.3.3, is a doubly stochastic process. This means that the parameters of an observed random process (ordnance deposition) are governed by another random process (target placement). Both sources of randomness must be taken into account in estimation to avoid underestimating uncertainty.

Our estimation methodology connects to the Neyman-Scott process through a result by Bartlet (1964), who has shown that the Neyman-Scott process is a Cox process (Cox, 1955) for which mathematically tractable constructions exist. The Cox process is based on the following postulates (Diggle 2003):

1. A spatial stochastic process generates a non-negative intensity function.
2. An inhomogeneous Poisson process produces events according to the generated intensity function.

A relatively flexible and tractable construction for a Cox process is the class of log-Gaussian processes (Moller et al., 1998). These models are tractable to the extent that spatial intensity estimation procedures are available via Markov chain Monte Carlo (MCMC).

Christensen and Waagepetersen (2002) use a log-Gaussian (LG) process model for estimating weed counts in a field from a sample of area counts for use in precision applications of weed killer. The analogy of applying weed killer only where there are weeds carries over to applying ordnance contamination remediation methods only where there are items of ordnance. Estimation of this model is implemented in a package geoRglm (Christensen and Ribeiro 2003) for the R statistical system (Ihaka and Gentleman 1996), which is distributed under the GNU General Public License.

The LG model assumes that the intensity is an exponential function of a Gaussian random component with a spatial correlation structure and a regression on a vector of covariates. In particular, for any location x within a site and a small region with area A centered on x, we assume that the number of ordnance objects in this region, given by $Y(x)$, has the Poisson distribution with a mean $\lambda(x)$. For each location x, we let

$$\lambda(x) = \exp(\, S(x) + f(x,\beta)\,),$$

where $S(x)$ is a correlated zero-mean Gaussian process, and $f(x,\beta)$ is a linear regression function on the initial, ASR-based, OIM upper and lower bounds in the neighborhood of x. Because of the Poisson distribution assumption, prediction equations do not have a closed form so that geoRglm proceeds by Markov chain Monte Carlo (MCMC, also sometimes known as the Metropolis-Hastings algorithm). Loosely speaking, we apply kriging to the log of the intensity, while treating the count data as the result of Poisson realizations.

As is customary in variations on kriging, this also requires the estimate of a variogram for the correlation function of the stochastic process $S(x)$. The package geoRglm provides two ways of estimating the variogram. This can be estimated externally to the MCMC estimation by fitting a variogram function to the log of the count data, or a Bayesian form can be estimated during the MCMC estimation.
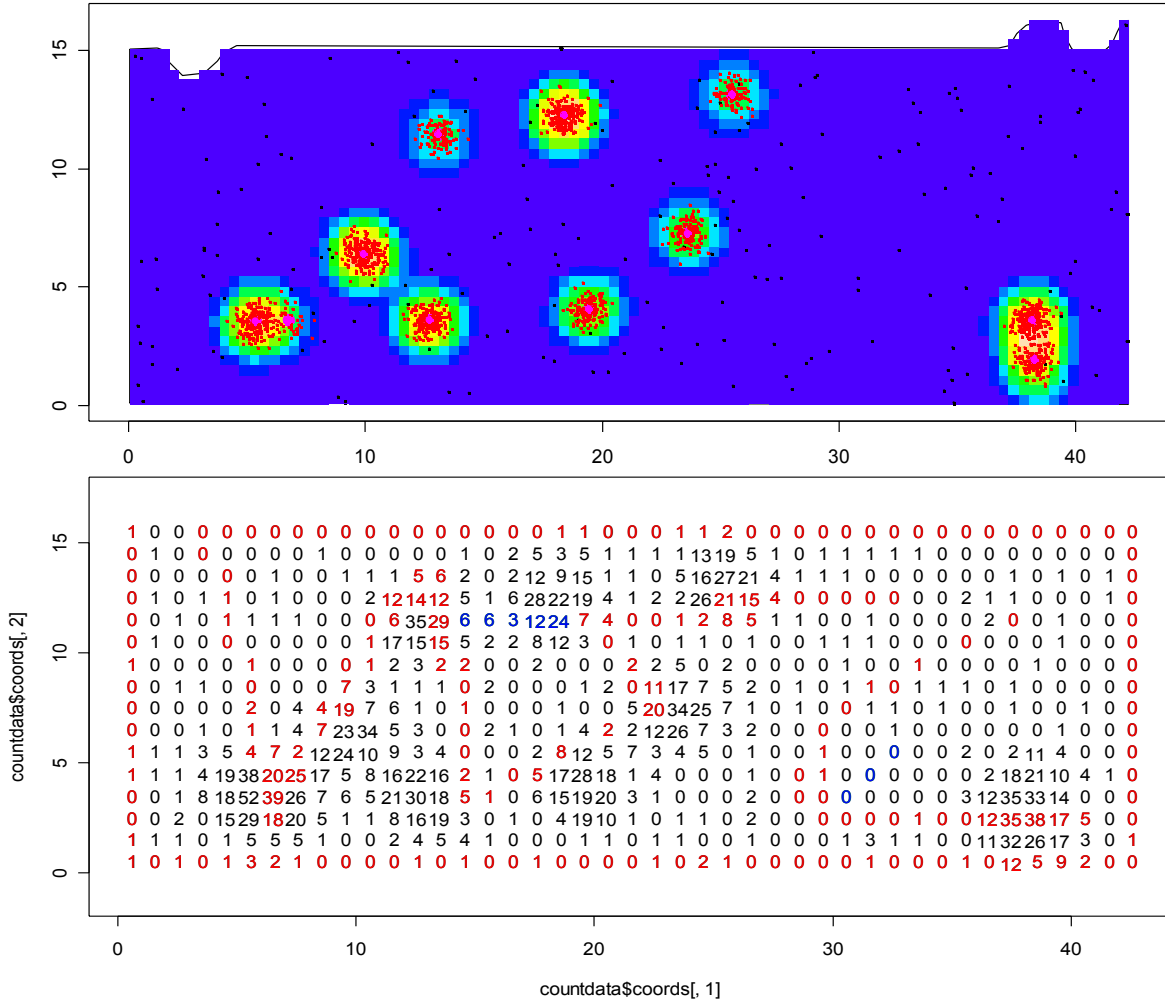
Fig. 37. Top: Simulated targets (magenta) with scattered ordnance (red) and background ordnance (black) positioned on a map of Badlands Bombing Range. Map color indicates ordnance intensity. Bottom: Ordnance counts after gridding top map. Red counts indicate locations to be sampled. Blue counts indicate locations predicted in Figs. 4 and 5.

We conducted a simulation with Gauss.target and generated several targets and associated ordnance on a Badlands Bombing Range map as illustrated in the top Figure 37. A gridding and counting function (that we include in Gauss.target) produces the counts in the bottom Fig. 37. Next, we consider a sample of this map indicated by the red counts. This sample alone is used in fitting a variogram and using geoRglm to produce distributions for all grids on the map, including those not sampled. Distribution histograms for the predicted intensity of several grid locations are in Figs. 38 and 39. Note that all true counts are covered by the predicted probability distributions (including the extreme 24 in Fig. 39, that is at about the .99 tail probability), confirming appropriate uncertainty coverage. Note that intensity estimates for the sampled grids also contain uncertainty. For example, a grid with a 1 count in the path on the left boundary of Fig. 39, results in an intensity estimate roughly between 0 and 3, as indicated by the green histogram. This is because of the Poisson stochastic layer of the underlying model; the

interpretation is that a Poisson distribution with intensity between 0 and 3 is likely to produce a count of 1 (although spatial correlation plays a role here too). The results we present in Figs. 4 and 5 are for the same simulated BBR map but with a different sample path.
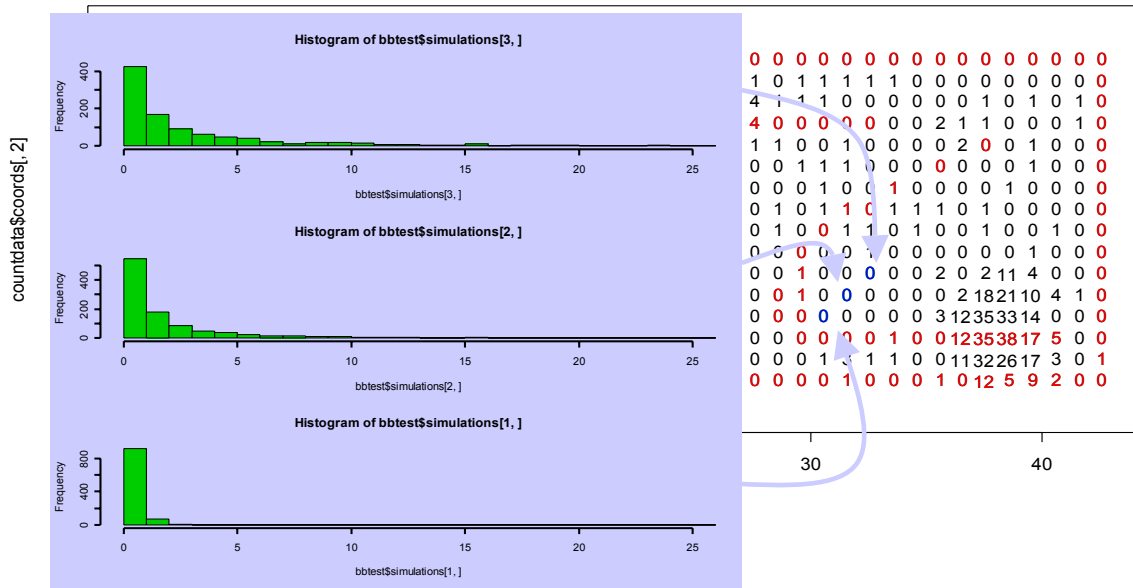
Fig. 38. Estimates of grids with true zero counts put more probability on bigger counts as distance from path increases. This is reasonable because correlation decreases with distance.
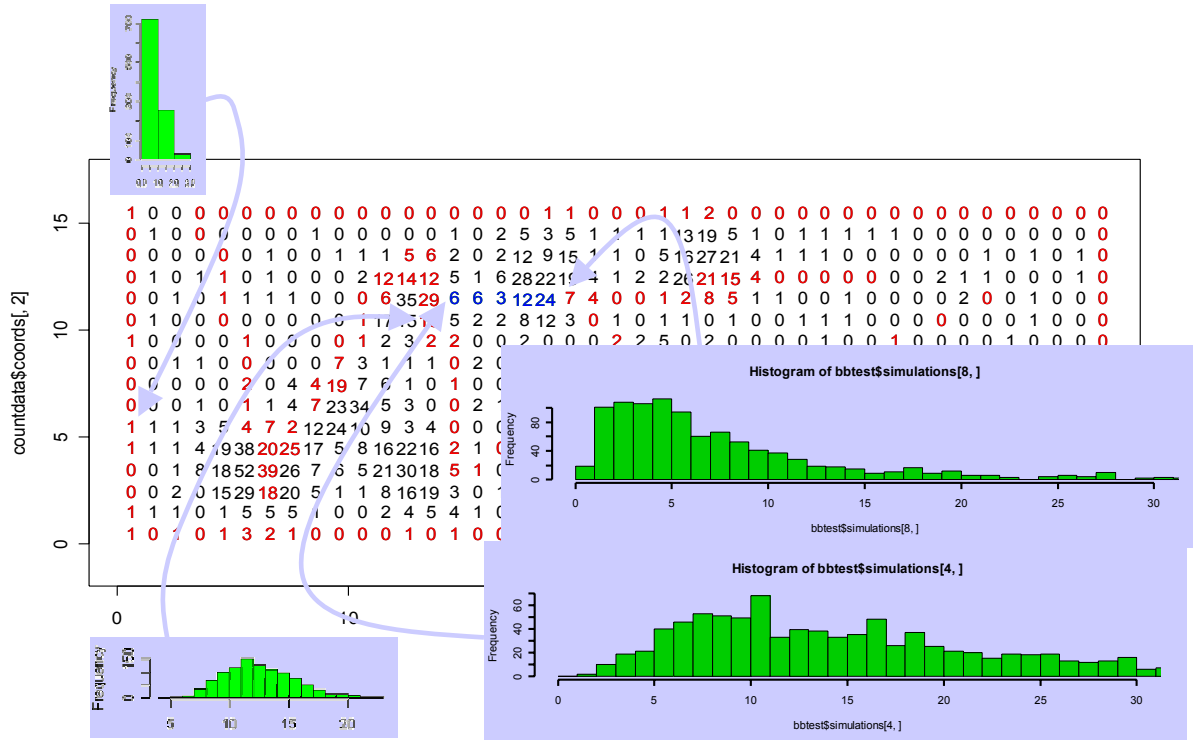
Fig. 39. Intensity distribution predictions for a number of grid locations. Note that uncertainty about intensity exists even at the grids that were sampled (top left histogram for the 1 and bottom left histogram for the 15).

2.6.2 Relationship to Other Kriging Methodologies

Count data sampled in portions of a grid can be interpolated with various kriging methodologies. The interpolation and variance calculations of simple kriging depend only on the sample locations and make no distributional assumptions other than spatial stationarity of the process. However, as soon as the computed variances are used in making a probability statement, an implicit distributional assumption on counts is made. Usually the assumption is Gaussian. The greatest difficulty with this assumption arises when the counts are low and the count distribution becomes asymmetric as is apparent in our estimates shown by histograms in Figs. 38 and 39 (in contrast to Gaussian symmetry). Using notation of the previous section, ordinary kriging uses the model

$$Y(x) = S(x),$$

dropping the Poisson layer of variability. In the same setup, the universal kriging model is

$$Y(x) = S(x) + f(x,\beta),$$

and the same implicit Gaussian assumption and loss of Poisson variability exists. The net result is an incorrect assessment of uncertainty.

This asymmetry in low count situations can be more appropriately handled by trans-Gaussian kriging. This methodology begins by taking a transformation of the data, so that a Gaussian assumption is more tenable, and proceeds with kriging. The results are then transformed with the inverse transformation. This introduces some bias (Cressie 1991) but can work well in some cases. In the case of log-Gaussian kriging, the model is

$$Y(x) = exp(S(x)).$$

However, both of the above kriging approaches are ignoring one stochastic layer that is present in our point process situation. They both ignore the Poisson nature of our count data. That is, they model $Y(x)$ instead of $\lambda(x)$. As a result, the uncertainty estimates will be overly optimistic.

One method that avoids the implicit Gaussian assumption on the counts and includes both stochastic layers is indicator kriging. Indicator kriging applies kriging to indicator functions of the data. For example, the function can specify if a threshold is exceeded with a certain minimum probability, where the probability computation can include both stochastic layers. Kriging is then applied to the probability that the indicator is 1. The reports on indicator kriging in the statistical literature are mixed. Cressie (1993) considers the theoretical underpinnings of indicator kriging to have some ad hoc components. Krivoruchko (2001) reports loss of information due to the indicator transformation. In any case, it seems clear that a careful evaluation and review of the theoretical properties of indicator kriging, and careful consideration of the method's appropriateness relative to the physical processes involved in the UXO mapping problem, is needed before it can be seriously considered in this context.

The advantage of all these kriging methods is that the estimation process has a closed form and can be computed very quickly. Our choice of the log-Gaussian model with MCMC estimation of Section 2.6.1, is based on its realistic assumptions, good theoretical underpinning, close relationship with the physics of ordnance deposition, and

the ability to estimate intensity and its uncertainty directly. As we have shown in our simulation, the estimation performs well for both low and high ordnance counts. Complete computation of our estimated DOAM distributions for the maps presented in Figs. 4 and 5 required about 20 minutes on an 850 MHz laptop computer. Subsequent production of all the maps from one DOAM representation takes only seconds. As current (2003) laptops are roughly 2 GHz and their speed doubles about every 18 months, we believe the computational requirements for DOAM are very reasonable.

2.7 Target Extent Delineation and Target Center Location

As we discussed in Section 1.1, the physical ordnance deposition process is a series of two-stage activities, where the first stage locates a target and the second stage scatters the ordnance around the target. Given a DOAM OIM representation of a few targets, here we discuss how to classify clean and contaminated areas on the basis of the DOAM OIM, and how to develop a DOAM TIM.

2.7.1 Target Extent Delineation

The DOAM OIM representation can be computed for any gridding resolution that is feasible on available computing resources. The resolution used in the BBR figures, such as Fig. 4, in this report is 16 x 43. Because all ordnance data are provided as locations, a single target or a few targets that are in close proximity can be estimated in greater detail with a finer resolution. Care needs to be taken to include a data border around the area of interest that is a little wider than the variogram. This is to include any data that may influence a prediction through spatial correlation. An estimated DOAM OIM representation can then provide a number of detailed maps useful for guiding further surveys necessary to completely classify an area according to some decision criterion.

Most regulatory situations require a threshold to be satisfied with some high probability. For example, if the threshold were 8 or fewer ordnance per grid with probability 0.90 or more, the blue areas in Fig. 5 would be considered clean. At the same time, grids in green, yellow, and orange in Fig. 4, have probability 0.20 or more that they have 10 or more ordnance per grid, so they could be considered as requiring remediation. Grids that satisfy neither of these requirements are candidates for further surveys until all grids are classified. These and similar maps at a greater resolution for a smaller area of the BBR can be used to classify areas according to some regulatory or remediation technology requirement.

So far, we have not addressed the difference between a UXO and exploded ordnance fragments because the focus of this report is the location and delineation of targets. However, we put forward a hypothesis that targets may typically have a UXO-free buffer around their perimeter that contains only exploded fragments. When ordnance explodes, it distributes fragments in a pattern that can be statistically defined. We have configured our Gauss.target simulation tool to provide this third level of ordnance deposition (1. locate target, 2. scatter ordnance hits around target, 3. scatter ordnance fragments around each ordnance hit). Figure 40 illustrates a simulated target that shows ordnance impact points and fragment scatter around those points. Both impact points and fragments are shown with Gaussian scatter. Because a UXO item remains at the impact point, this suggests that UXO scatter is typically less than fragment scatter and

targets may have a UXO-free border. Depending on the explosion scatter radius, the UXO-free border may cover a substantial portion of target area as illustrated in Fig. 40. We suggest that this is checked by ground truth data from a few targets.
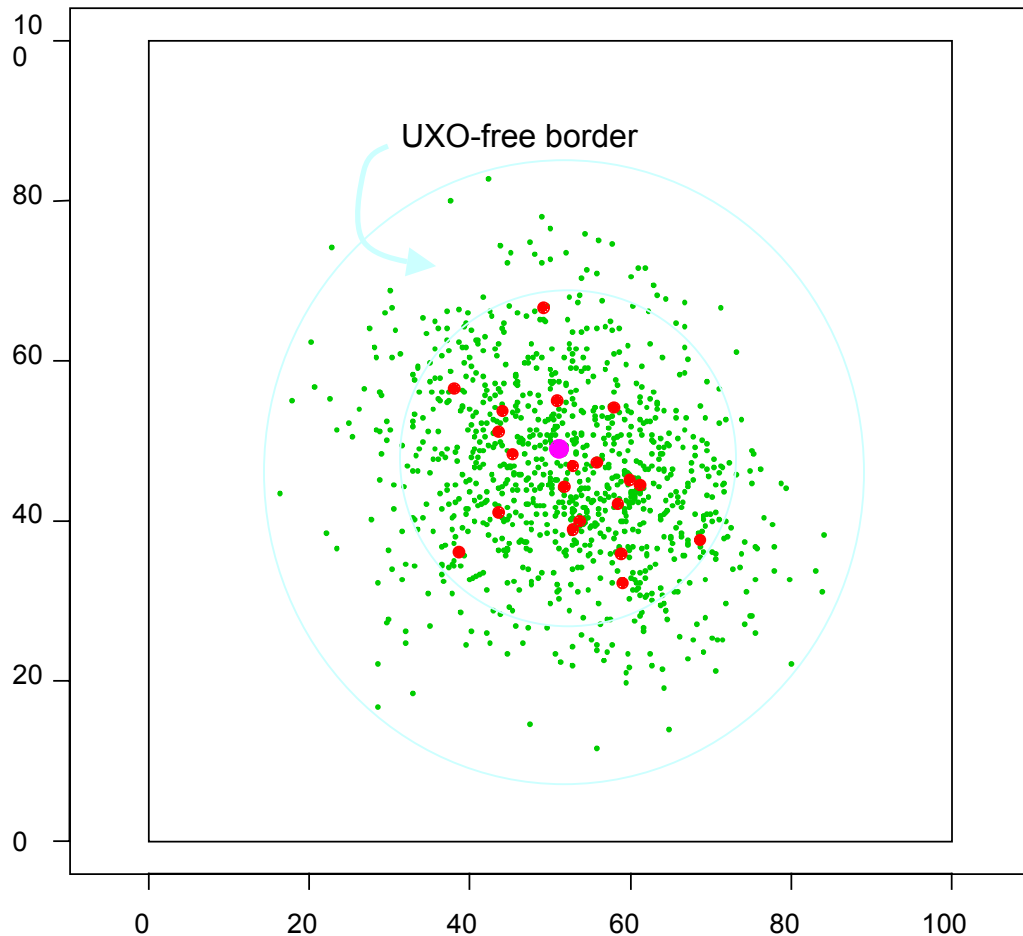
Fig. 40. Gauss.target simulation: Target center (magenta) with Gaussian scatter of ordnance impact points (red) and Gaussian ordnance fragment scatter around each impact point (green).

2.7.2 Target Center Location and TIM estimation

Ordnance locations are a *consequence* of target location, therefore, it must be information other than ordnance location that influences target location. The location of targets may be influenced by site configuration such as topography, vegetation, geology, training objectives, logistics, safety, and the location of other targets or structures on the site. As a result, it makes sense to talk about a point process describing target locations

as a model of human decisions to locate targets within a site. This is an important concept that was not envisioned at the outset of this project and consequently we only describe it without providing software or estimates. We believe that this could become a very effective means of locating undocumented targets within a large site by using information from all targets across all sites.

A procedure similar to that of ordnance location from a geophysical signal can be used to locate targets in a map of site ordnance intensity. The resulting target locations can be converted to grid counts on a rather coarse spatial resolution. These will be low counts, so an appropriate model is a log-Gaussian point process with an intensity

$$\lambda(x) = \exp(\ S(x) + f(x,\beta)\ ),$$

where $S(x)$ is a correlated zero-mean Gaussian process, and $f(x,\beta)$ is a linear regression function on a set of grid features. The grid features should include anything that may influence a human decision to locate a target. For example, relative elevation, gradient, second order gradient, type of vegetation, distance to nearest occupied structure, prevailing wind direction, etc. This model can be estimated with geoRglm and predictions of target location intensity in the form of a DOAM TIM can be made for an entire site. The TIM can then be used to compute maps such as the probability that one or more targets are located in a grid.

3 The Site Characterization Process Flow

The complete site characterization process that leads to and maintains the DOAM OIM and TIM representations is presented in Fig. 41. It repeats the diagram presented at the outset of Section 2, while providing some additional detail. Some of the components in Figure 4.1 are conceptual and need further development. Some require implementation on a site-specific basis, and some have production or research software codes available Further projects need to be initiated to complete these tasks before an integrated UXO remediation suite of software tools can be produced and routinely used.
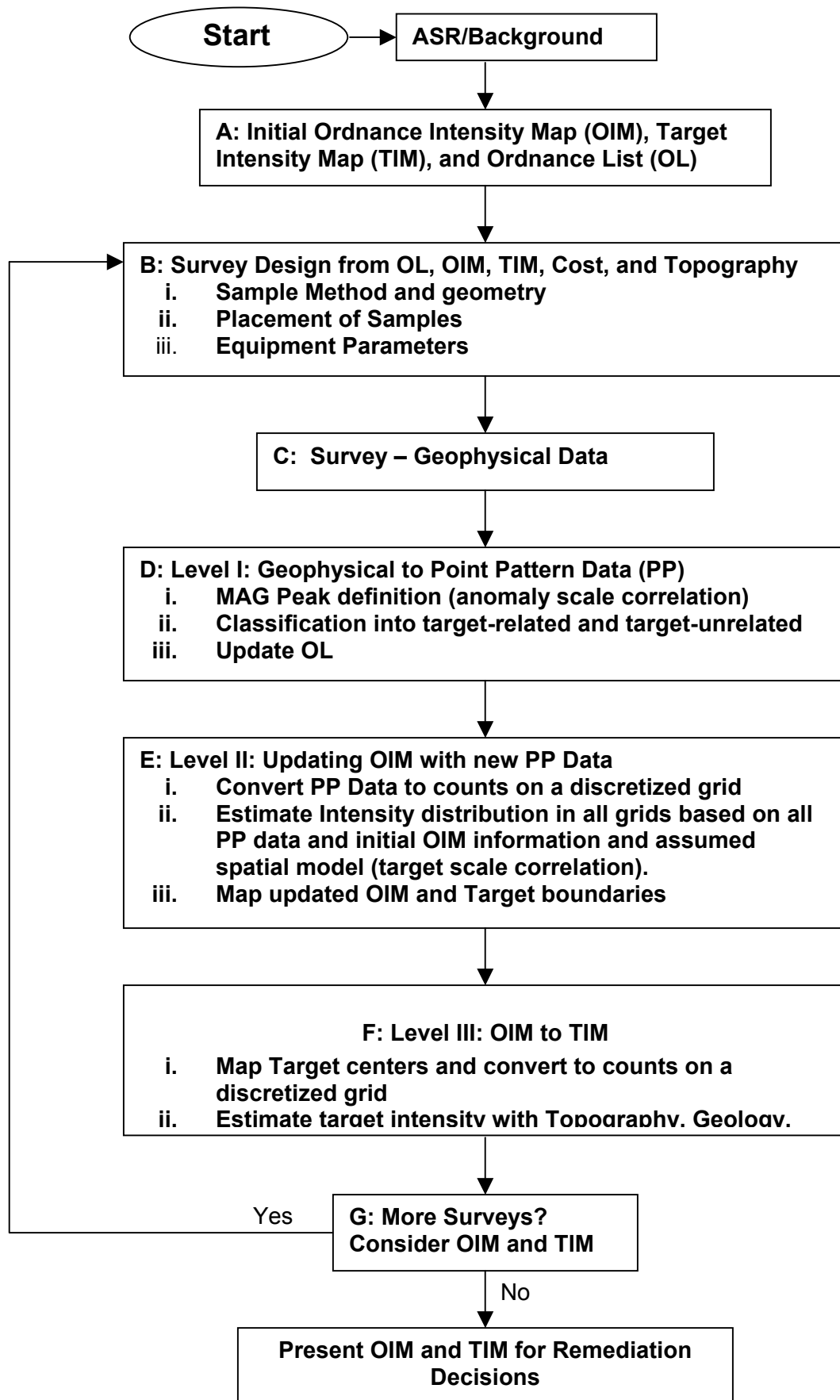
Fig. 41. Flowchart for statistical design and analysis of UXO surveys.

REFERENCES

Adley, M. A., Berger, R. P., and Cargile, J. D., 2003. Methodology and User's Guide for PENCURV +. Technical Report ERDC/GSL TR-02-2. U.S. Army Engineer Research and Development Center, Vicksburg, MS.

Altshuler, T. W., 1996. Shape and Orientation Effects on Magnetic Signature Prediction for Unexploded Ordnance Proceedings of the UXO Forum 1996. pp. 282291, Williamsburg, VA.

Asch, T., Cormier, M. J., Klaff, T. L., Murray, C., Petersen, G., Staes, E. G. 2002. Receiver Operating Characteristics Curves in the Ordnance Detection and Discrimination Study at the Former Fort Ord, California. 12 pp. Proceedings, SAGEEP 2002. Available on CD-ROM from Environmental and Engineering Geophysical Society, Denver, CO.

Barrett, B., Fanning, A. 1999. UXO Calculator: a new statistical approach for determining unexploded ordnance (UXO) density at ordnance sites. UXO Forum 1999. Atlanta, GA, May 25—27, 1999.
http://www.uxocoe.brtrc.com/UXOForumDocs/Forum99/DT_Fanning.pdf

Billings, S. D., Stanley, J. M., and Youmans, C. 2002. Magnetic Discrimination that will Satisfy Regulators. Proceedings of the UXO/Countermine Forum 2002 (CD-ROM). Orlando, FL.

Bilisoly, R. L. and McKenna, S. A. 2003. Determining Optimal Location and Numbers of Sample Transects for Characterization of UXO Sites. SAND Report SAND2002-3962. 47 pgs. January. 2003. Sandia National Laboratories, Albuquerque, NM.

Butler, D. K., Cespedes, E. R., Cox, C. B., and Wolfe, P. J. 1998. Multisensor Methods for Buried Unexploded Ordnance Detection, Discrimination and Identification, SERDP-98-10. U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Butler, D. K., Wolfe, P. J., and Hansen, R. O. 2001. Analytical Modeling of Gravity and Magnetic Signatures of Unexploded Ordnance. Journal of Environmental and Engineering Geophysics. Vol. 6, No. 1. pp. 33-46.

Butler, D. K., Pasion, . R., Billings, S., Oldenburg, D., and Yule, D. 2003. Model-based Inversion for Enhanced UXO Detection and Discrimination. Proceedings of the Detection of Mines and Mine-like Targets Conference. Society of Profession Optical Engineers (SPIE), Orlando, FL. AeroSense 2003.

Byers, S. and Raftery, A. E. 1998. Nearest-Neighbor Clutter Removal for Estimating Features of Spatial Point Processes. Journal of the American Statistical Association 93:577-584.

Christensen, Ole F. and Ribeiro Jr., Paulo J. 2003. geoRGlm, Software for generalised linear spatial models using R.
http://www.maths.lancs.ac.uk/~christen/geoRglm/

Christensen, O. F. and Waagepetersen, R. P. 2002. Bayesian prediction of spatial count data using generalised linear mixed models. Biometrics *58*:280–286.

Cox, D.R. 1955. Some Statistical Methods Connected with Series of Events. Journal of the Royal Statistical Society B. 17:29-164.

Cressie, N. A. C. 1993. Statistics for Spatial Data. Wiley, New York.

Department of the Army. 1986. Fundamentals of Protective Design for Conventional Weapons. Technical Manual TM 5-855-1. Washington, DC.

Department of the Army. 2000. Engineering and Design—Ordnance and Explosives Response. Engineer Manual EM 1110-1-4009. 23 June 2000, Washington, DC.

Diggle, P. 2003. Statistical Analysis of Spatial Point Patterns (second edition). Oxford University Press.

Diggle, P., Tawn, J. A., and Moyeed, R. A. 1998. Model-based geostatistics (with discussion). Applied Statistics. 47:299-350.

Doll, W. E., Gamey, T. J., and Holladay, J. S. 2001. Current Research into Airborne UXO Detection. Extended abstract. Proceedings of the 2001 Symposium on the Application of Geophysics to Environmental and Engineering Problems. 16pp. Available on CD-ROM.

Doll, W.E., Gamey, T. J., Beard, L. P., Bell, D. T., and Holladay, J. S. 2003. Recent advances in airborne survey technology yield performance approaching ground-based surveys. The Leading Edge. 22:420-425.

Geosoft. 1997. Oasis Montaj Data Processing and Analysis System for Earth Science Applications, Version 4.1 User Guide. 290 pp. Chapter 7: Processing Data.

Geosoft. 2003. Oasis Montaj is a product of Geosoft Inc. Information can be found online at http://www.geosoft.com .

Gilbert, R. O., Wilson, J. E., O'Brien, R. F., Carlson, D. K., Bates, D. J., Pulsipher, B. A., McKinstry, C. A. 2002. Version 2.0 Visual Sample Plan (VSP): Models and Code Verification. PNNL Report PNNL-13991. August 2002. 108 pages. Pacific Northwest National Laboratory, Richland, WA. (http://dqo.pnl.gov/VSP/pnnl13991.pdf)

Gnanadesikan, R. 1977. Methods for Statistical Data Analysis of Multivariate Observations. John Wiley & Sons, New York.

Hassig, N. L., O'Brien, R. F., Wilson, J. E., Pulsipher, B. A., Gilbert, R. O., McKinstry, C. A., Carlson, D. K., Bates, D. J. 2002. Visual Sample Plan Version 2.0 User's Guide. PNNL Report PNNL-14002. September 2002. 95 pages Pacific Northwest National Laboratory, Richland, WA. (http://dqo.pnl.gov/VSP/PNNL14002.pdf)

Hogg, R.V and Craig, A. T. 1970. Introduction to Mathematical Statistics. Macmillan Publishing Co., Inc. New York.

Ihaka, R., Gentleman, R. 1996. R: A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics. 5:299–314.

Krivoruchko, K. 2001. Using linear and non-linear kriging interpolators to produce probability maps. 2001 Annual Conference of the International Association for Mathematical Geology. Cancun, Mexico. http://www.kgs.ukans.edu/Conferences/IAMG/Sessions/K/Papers/krivoruchko2.pdf

McFee, J. E. 1989. "Electromagnetic Remote Sensing: Low Frequency Electromagnetics." Technical Report 124. Defence Research Establishment Suffield, Ralsotn, Alberta, CA.

Moller, J., Syversveen, A. R., Waagepetersen, R. P. 1998. Log Gaussian Cox Processes. Scandinavian Journal of Statistics. 25:451–482.

McDonald, J., Nelson, H., Neece, J., Robertson, R. Jeffries, R. 1998. MTADS Unexploded Ordnance Operations at the Badlands Bombing Range on the Pine Ridge Reservation, Cuny Table, SD. NRL/PU/6110–98-353.

McDonald, J. R., Robertson, R. 1996. Sensor Evaluation Study for Use With Towed Arrays for UXO Site Characterization. Proceedings of SAGEEP '96. pp. 451–464.

Nelson, H. H., Altshuler, T. W., Rosen, E. M., McDonald, J. R., Barrow, B., Khadr, N. 1998. Magnetic Modeling of UXO and UXO-like Targets and Comparison with Signatures Measured by MTADS. UXO Forum 1998. Anaheim, CA. May 28–30, 1998. Conference Proceedings: (http://www.uxocoe.brtrc.com/UXOForumDocs/Forum98/Nelson.pdf).

Neyman, J., Scott, E. L. 1958. Statistical approach to problems of cosmology, *Journal of the Royal Statistical Society, Series B*. 20 (1):1–43.

Neyman, J., Scott, E. L. 1972. Processes of clustering and applications. *Stochastic Point Processes, Statistical Analysis, Theory and Applications*. P. A. W. Lewis, ed. Wiley, New York. pp. 646–681.

ORNL. 2000. Evaluation of Footprint Reduction Methodology at the Cuny Table in the Former Badlands Bombing Range, Project Final Report to Environmental Security Technology Certification Program (ESTCP). 37 pp. and 6 appendices with a CD-ROM.

Ostrouchov, G., G. P. Zimmerman, J. J. Beauchamp, V. V. Federov, Downing, D. J. 1999. Evaluation of statistical methodologies used in U.S. Army ordnance and explosive work. ORNL/TM-13588. Oak Ridge National Laboratory, Oak Ridge TN. September 1999. 34 pp.

Panel on Discriminant Analysis, Classification, and Clustering. 1989. Discriminant Analysis and Clustering. *Statistical Science*, **4**. pp 34–69.

Pepe, M. S. 2000. Receiver Operating Characteristic Methodology. *Journal of the American Statistical Association*, **95**. pp. 308–311.

Quantitech. 1995. Ordnance and explosives site statistical sampling based methodology (SiteStats) final report. Technical report 95-R-011, QuantiTech, Huntsville, AL. September 30, 1995.

Quantitech. 1995. Grid statistical sampling based methodology (GridStats), Version 1.2, User's Manual. Huntsville, AL.   Prepared for U.S. Army Engineer Division. Huntsville, AL. September 30, 1995.

Ripley, B. R. 1988. *Statistical Inference for Spatial Processes*, Cambridge University Press. Cambridge.

Stoyan, D., Kendall, W. S., Mecke, J. 1995. *Stochastic Geometry and its Applications, Second Edition*. John Wiley & Sons, New York.

Tarno, L., Butler, J. 1986. Range Clearance Technology Assessment. Technical Report NAVEODTECHCEN TR-275. Indian Head, MD.

USAESCH. 1999. Draft. Archive Search Report Findings for the former Pine Ridge Gunnery Range. Shannon and Jackson Counties, South Dakota. Project Number B08SD001802. April 1999.

USAESCH. 2003. Online article entitled "ASR Process." Found at http://www.mvs.usace.army.mil/engr/ed-p/asr.htm

Appendices

Appendix 1: R files on Enclosed CD-ROM

Directory Listing:

```
07/08/2003  10:17p          2,465          gauss.targets.r.txt
07/24/2003  12:41p          9,090          Splus-SegmentExchange.txt
01/13/2003  02:53p      8,345,513          Rtarget.RData
01/13/2003  02:53p          4,571 .        Rhistory
07/24/2003  01:11p     22,257,262          rw1071.exe
07/24/2003  01:11p         11,031          README.rw1071
07/24/2003  01:12p        713,408          geoR.zip
07/24/2003  01:13p        313,318          geoRglm.zip
07/24/2003  01:13p        353,025          splancs.zip
07/24/2003  01:17p        204,157          mva_code.tar.gz
              10 File(s)    32,213,840 bytes
```

gauss.targets.r.txt - Gauss.target simulation tool R-code. Can be imported into an R session with source() command.

Splus-SegmentExchange.txt - S-plus code for Segment Exchange algorithm, may require Splus package, although is likely to run under R also.

Rtarget.RData - An R workspace that contains all functions for DOAM estimation as well as supporting codes for Gauss.target. This is the first thing to load into R, followed by splancs, geoR, and geoRglm below.

.Rhistory - contains a history of recent commands in an R session and may be useful as examples of how to run various functions.

rw1071.exe - an executable for R setup on a Windows platform (linux and Mac versions are available from www.r-project.org)

README.rw1071 - installation instructions for R

geoR.zip - a general geostatistics package for R

geoRglm.zip - the R package for estimating DOAM

splancs.zip - an R package required for Gauss.target

mva_code.tar.gz - R code for multivariate analysis of target related and target unrelated anomalies

Appendix 2 Summary Statistics for magnetic survey parameters, overall and by anomaly group

Data from Badlands Bombing Range, South Dakota.

```
Parameter: Offset angle of total field anomaly from magnetic north (deg) [theta]
          n n=0  Mean   SD   Min  0.05   0.1  0.25   0.5 0.75  0.9 0.95   Max
Overall 345   0  4.15 40.4 -89.5 -64.5 -48.6 -24.0   5.9 31.8 58.9 72.5 89.7
 - ord  180   0  6.36 38.7 -89.5 -63.2 -41.4 -19.3  10.0 30.9 55.4 69.1 84.1
 - dirt  27   0 -5.54 55.9 -86.0 -79.6 -62.6 -57.2 -15.2 40.3 73.1 78.0 87.1
 - scrap 63   0  5.18 40.0 -80.0 -53.8 -40.1 -23.3   0.5 35.8 67.9 72.7 89.7
 - wire  75   0  1.45 38.6 -80.0 -65.8 -55.3 -24.1   4.3 25.9 53.2 61.4 88.1


Parameter: Analytic signal (nT/m) [as.grid]
          n n=0  Mean    SD Min 0.05  0.1 0.25   0.5 0.75   0.9  0.95    Max
Overall 345   0 16.60 29.90 1.3 2.70 3.94 5.70  8.10 13.9  29.5  56.3  320.0
 - ord  180   0  9.97  7.75 1.6 4.30 4.80 5.70  7.55 12.0  16.5  21.0   67.3
 - dirt  27   0  9.45 12.80 1.3 1.64 2.26 3.85  6.30  9.2  15.3  23.9   67.5
 - scrap 63   0 33.20 58.20 1.9 2.95 3.62 6.45 11.20 24.4 106.0 140.0  320.0
 - wire  75   0 21.20 27.20 1.6 2.24 3.06 6.00 10.30 25.8  46.9  77.0  132.0


Parameter: Magnetic peak value (nT) [magpeak.value]
          n n=0 Mean    SD  Min  0.05   0.1  0.25  0.5 0.75    0.9   0.95 Max
Overall 345   0 44.3  81.7 -9.6  6.60  9.70 15.0 21.6 36.7   83.3 147.0 826
 - ord  180   0 26.5  19.7  3.6 12.10 13.50 16.2 21.1 29.1   47.4  57.7 196
 - dirt  27   0 25.3  30.2  1.8  5.07  6.52 10.6 13.9 30.0   49.1  75.3 147
 - scrap 63   0 92.4 164.0  5.9  8.52  9.08 14.9 27.0 78.9  251.0 349.0 826
 - wire  75   0 53.7  65.7 -9.6  4.44  6.44 12.5 28.1 73.0  139.0 212.0 325


Parameter: Magnetic low value (nT) [maglow.value]
          n n=0   Mean   SD  Min   0.05   0.1  0.25    0.5  0.75   0.9  0.95  Max
Overall 345   1 -10.10 20.8 -168 -40.6 -25.70 -10.0  -3.40 -1.70 -0.30  0.64 11.5
 - ord  180   1  -4.06 11.0 -140 -11.4  -6.64  -4.4  -2.55 -1.20 -0.19  0.71  6.6
 - dirt  27   0 -16.50 17.0  -59 -56.1 -40.60 -21.7 -12.00 -3.80 -2.74 -0.78  3.7
 - scrap 63   0 -18.00 31.5 -165 -82.5 -42.20 -19.0  -5.90 -1.85 -0.64  3.33 11.5
 - wire  75   0 -15.50 24.5 -168 -56.7 -37.30 -18.5  -6.60 -2.85 -1.58 -0.65  1.3


Parameter: Magnetic peak-to-peak amplitude (nT) [mag.pp]
          n n=0  Mean    SD  Min  0.05  0.1 0.25  0.5 0.75  0.9  0.95 Max
Overall 345   0  54.4  91.6 4.80 10.50 14.5 19.3 27.3  47.9 113 213.0 870
 - ord  180   0  30.6  25.4 7.01 14.30 15.9 19.2 23.0  32.4  51  67.6 227
 - dirt  27   0  41.8  38.3 4.80  6.89 10.7 15.8 33.4  47.3 103 134.0 143
 - scrap 63   0 110.0 178.0 7.04 10.90 14.4 20.6 38.0 101.0 303 377.0 870
 - wire  75   0  69.2  77.7 5.85  8.84 10.7 19.3 36.7  84.2 177 252.0 366


Parameter: Magnetic peak separation (m) [mag.pp.sep]
          n n=0 Mean   SD  Min 0.05  0.1 0.25  0.5 0.75  0.9 0.95  Max
Overall 345   0 13.0 4.54 6.18 7.50 8.08 10.0 12.2 15.4 18.5 20.6 35.3
 - ord  180   0 13.8 4.12 6.18 7.50 8.98 11.2 12.9 16.1 19.6 20.9 29.8
 - dirt  27   0 12.4 5.46 6.71 7.55 7.91  9.6 12.1 13.4 17.2 18.2 35.2
 - scrap 63   0 12.6 5.16 6.18 7.50 7.53  9.0 11.3 15.3 17.8 22.6 30.4
 - wire  75   0 11.8 4.33 6.36 7.50 7.65  9.0 10.6 13.9 17.1 18.5 35.3


Parameter: Separation of magnetic peak and analytic signal (m) [magpeak.as.sep]
          n n=0 Mean   SD Min 0.05 0.1 0.25  0.5 0.75  0.9  0.95  Max
Overall 345  76 2.12 3.03   0 0.00 0.0  0.5 1.50 2.12 5.13  8.72 19.5
 - ord  180  42 1.31 1.46   0 0.00 0.0  0.5 1.12 1.58 2.24  3.17 10.6
 - dirt  27   2 4.06 4.03   0 0.45 1.5  1.5 2.12 5.37 8.69 12.60 16.2
 - scrap 63  12 2.75 3.28   0 0.00 0.0  0.5 1.50 3.18 7.82  9.31 15.9
 - wire  75  20 2.86 4.35   0 0.00 0.0  0.0 1.50 2.12 8.31 13.80 19.5


Parameter: Width of analytic signal (m) [width.as]
          n n=0 Mean    SD Min 0.05  0.1 0.25 0.5 0.75  0.9 0.95  Max
Overall 345   0 3.11 1.560 1.8  1.8 1.80 2.00 2.7 3.70 4.86 6.08 11.1
 - ord  180   0 3.01 1.340 1.8  1.8 1.80 2.00 2.7 3.60 4.51 5.41  9.6
 - dirt  27   0 2.85 0.873 1.8  1.8 1.86 2.15 2.8 3.25 3.88 4.63  4.9
 - scrap 63   0 3.41 2.000 1.8  1.8 1.80 2.00 2.6 4.40 6.32 7.36 11.1
```

Parameter: (continued)

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
- wire	75	0	3.20	1.780	1.8	1.8	1.80	1.90	2.7	3.90	4.76	8.03	10.3

Parameter: Width of total field peak (m) [width.tf.peak]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	0	2.55	1.18	1.4	1.8	1.8	1.90	2.4	2.50	3.30	5.06	10.5
- ord	180	0	2.58	1.26	1.4	1.8	1.8	2.00	2.4	2.42	3.01	5.10	10.5
- dirt	27	0	2.34	0.62	1.8	1.8	1.8	1.85	2.2	2.50	3.18	3.30	4.4
- scrap	63	0	2.60	1.14	1.8	1.8	1.8	1.90	2.4	2.55	3.54	5.40	7.8
- wire	75	0	2.50	1.17	1.8	1.8	1.8	1.80	2.4	2.40	3.30	3.93	9.1

Parameter: Ratio, magpeak.value : mag.pp [rat.mpv.mpp]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	0	0.813	0.199	0.102	0.402	0.543	0.731	0.865	0.941	0.986	1.020	1.66
- ord	180	0	0.879	0.134	0.204	0.670	0.747	0.829	0.905	0.955	0.992	1.030	1.36
- dirt	27	0	0.595	0.208	0.192	0.289	0.329	0.434	0.595	0.724	0.872	0.919	1.03
- scrap	63	0	0.786	0.236	0.230	0.439	0.517	0.614	0.838	0.941	0.971	1.020	1.66
- wire	75	0	0.758	0.220	0.102	0.263	0.429	0.681	0.826	0.892	0.953	0.982	1.08

Parameter: Ratio, magpeak.as.sep : width.was [rat.mpass.was]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	76	0.786	1.130	0	0.00	0.000	0.208	0.500	0.833	1.760	3.02	8.34
- ord	180	42	0.482	0.473	0	0.00	0.000	0.171	0.409	0.697	0.941	1.22	3.31
- dirt	27	2	1.490	1.410	0	0.13	0.454	0.572	0.884	1.840	3.530	4.78	5.05
- scrap	63	12	1.000	1.250	0	0.00	0.000	0.236	0.625	1.330	2.560	4.22	5.73
- wire	75	20	1.080	1.690	0	0.00	0.000	0.000	0.589	0.949	3.460	4.32	8.34

Parameter: Ratio, magpeak.as.sep : width.tf.peak [rat.mpass.wtfp]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	76	0.895	1.310	0	0.00	0.000	0.179	0.600	0.833	2.010	4.06	8.39
- ord	180	42	0.542	0.667	0	0.00	0.000	0.142	0.457	0.750	0.972	1.21	5.05
- dirt	27	2	1.770	1.690	0	0.18	0.615	0.743	0.922	1.980	4.460	5.06	6.43
- scrap	63	12	1.170	1.520	0	0.00	0.000	0.208	0.707	1.250	3.570	4.05	8.39
- wire	75	20	1.200	1.800	0	0.00	0.000	0.000	0.625	0.833	4.020	5.28	7.53

Parameter: Ratio, magpeak.as.sep : ihpd [rat.mpass.ihpd]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	76	0.958	1.350	0	0.0000	0.000	0.196	0.528	0.993	3.110	4.55	6.86
- ord	180	42	0.525	0.673	0	0.0000	0.000	0.161	0.434	0.663	0.953	1.45	4.40
- dirt	27	2	2.110	1.860	0	0.0697	0.399	0.863	1.220	3.560	4.860	5.19	5.92
- scrap	63	12	1.360	1.580	0	0.0000	0.000	0.219	0.688	1.720	4.060	4.76	5.77
- wire	75	20	1.250	1.700	0	0.0000	0.000	0.000	0.645	1.190	4.470	4.72	6.86

Parameter: Ratio, magpeak.as.sep : mag.pp.sep [rat.mpass.mpps]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	76	0.180	0.253	0	0.0000	0.0000	0.0411	0.1010	0.175	0.557	0.857	1.380
- ord	180	42	0.104	0.134	0	0.0000	0.0000	0.0324	0.0873	0.127	0.192	0.252	0.884
- dirt	27	2	0.360	0.325	0	0.0128	0.0664	0.1490	0.2210	0.600	0.868	0.889	1.020
- scrap	63	12	0.249	0.294	0	0.0000	0.0000	0.0440	0.1320	0.295	0.791	0.877	1.000
- wire	75	20	0.239	0.332	0	0.0000	0.0000	0.0000	0.1240	0.210	0.848	0.924	1.380

Parameter: Ratio, mag.pp.sep : ihpd [rat.mpps.ihpd]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	0	5.30	0.458	4.72	4.96	4.97	4.97	4.98	5.45	6.21	6.22	6.25
- ord	180	0	5.12	0.257	4.72	4.96	4.96	4.97	4.98	5.43	5.44	5.45	6.24
- dirt	27	0	5.95	0.341	5.42	5.43	5.44	5.61	6.20	6.21	6.22	6.23	6.25
- scrap	63	0	5.44	0.525	4.96	4.96	4.96	4.97	5.42	6.00	6.21	6.22	6.25
- wire	75	0	5.39	0.534	4.96	4.96	4.97	4.97	4.98	5.99	6.21	6.22	6.25

Parameter: Ratio, width.as : ihpd [rat.was.ihpd]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	0	1.48	1.150	0.337	0.542	0.615	0.793	1.110	1.71	2.82	3.57	8.51
- ord	180	0	1.25	0.798	0.354	0.528	0.572	0.731	0.978	1.55	1.99	2.90	5.82
- dirt	27	0	1.53	0.708	0.495	0.716	0.929	1.100	1.260	1.78	2.69	2.96	3.25
- scrap	63	0	1.81	1.470	0.337	0.557	0.593	0.796	1.230	2.10	4.40	4.89	6.59
- wire	75	0	1.74	1.520	0.431	0.607	0.655	0.853	1.270	2.03	3.19	4.70	8.51

Parameter: Ratio, width.tf.peak : ihpd [rat.wtfp.ihpd]

	n	n=0	Mean	SD	Min	0.05	0.1	0.25	0.5	0.75	0.9	0.95	Max
Overall	345	0	1.15	0.758	0.308	0.544	0.636	0.802	0.978	1.24	1.61	2.19	7.24
- ord	180	0	1.06	0.734	0.329	0.544	0.606	0.741	0.891	1.09	1.54	2.13	7.24
- dirt	27	0	1.25	0.557	0.387	0.843	0.915	0.970	1.160	1.34	1.56	1.89	3.58
- scrap	63	0	1.27	0.799	0.473	0.528	0.586	0.867	1.130	1.46	1.70	2.18	5.91

```
 - wire    75    0 1.25 0.822 0.308 0.659 0.736 0.889 1.160 1.29 1.59 2.10 7.00


Parameter: Ratio, width.tf.peak : width.as [rat.wtfp.was]
           n n=0 Mean   SD   Min  0.05   0.1  0.25   0.5 0.75  0.9 0.95  Max
Overall  345    0 0.956 0.545 0.175 0.357 0.462 0.632 0.900 1.20 1.33 1.60 5.56
 - ord   180    0 0.994 0.641 0.188 0.404 0.475 0.645 0.879 1.22 1.33 1.63 5.56
 - dirt   27    0 0.867 0.236 0.408 0.554 0.581 0.670 0.909 1.05 1.16 1.19 1.32
 - scrap  63    0 0.935 0.482 0.257 0.316 0.357 0.523 0.947 1.27 1.49 1.68 2.85
 - wire   75    0 0.915 0.403 0.175 0.358 0.455 0.615 0.900 1.23 1.33 1.57 2.00


Parameter: Ratio, mag.pp.sep : width.as [rat.mpps.was]
           n n=0 Mean   SD   Min  0.05   0.1 0.25   0.5 0.75  0.9 0.95   Max
Overall  345    0 5.01 2.53 0.728 1.60 2.00 3.14 4.74 6.45 8.49 9.38 14.8
 - ord   180    0 5.33 2.52 0.938 1.71 2.50 3.37 5.16 6.88 8.79 9.87 14.1
 - dirt   27    0 4.68 2.11 1.840 1.93 2.08 3.29 4.61 5.26 6.74 8.72 11.0
 - scrap  63    0 4.76 2.83 0.879 1.23 1.39 2.64 4.42 6.24 8.56 8.94 14.8
 - wire   75    0 4.55 2.35 0.728 1.40 1.91 2.66 4.47 5.83 7.59 8.20 12.6


Parameter: Ratio, mag.pp.sep : width.tf.peak [rat.mpps.wtfp]
           n n=0 Mean   SD   Min  0.05  0.1 0.25   0.5 0.75  0.9 0.95   Max
Overall  345    0 5.58 2.18 0.687 2.50 3.44 4.25 5.27 6.72 8.16 9.32 17.7
 - ord   180    0 5.90 2.17 0.687 2.34 3.42 4.74 5.85 6.87 8.69 9.32 16.5
 - dirt   27    0 5.43 2.12 1.740 3.15 3.74 4.53 5.33 5.96 6.72 7.34 14.1
 - scrap  63    0 5.26 2.24 0.981 2.58 3.45 3.82 4.72 6.03 8.54 9.42 12.1
 - wire   75    0 5.13 2.08 0.824 2.57 3.57 4.17 4.86 5.83 7.13 7.60 17.7


Parameter: Calculated anomaly depth (m) [depth]
           n n=0  Mean   SD   Min   0.05    0.1  0.25    0.5 0.75  0.9 0.95   Max
Overall  345    0 0.670 0.986 -7.08 -0.526 -0.170 0.190 0.660 1.120 1.74 2.14 4.25
 - ord   180    0 0.854 0.985 -7.08 -0.340 -0.051 0.355 0.865 1.270 1.90 2.22 3.32
 - dirt   27    0 0.935 0.794 -0.07  0.081  0.350 0.540 0.810 1.070 1.43 1.85 4.25
 - scrap  63    0 0.372 1.220 -5.55 -1.000 -0.504 0.000 0.380 0.870 1.45 1.76 4.10
 - wire   75    0 0.383 0.663 -1.57 -0.470 -0.296 0.065 0.330 0.685 1.05 1.40 2.87


Parameter: Instrument height (m) [inst.height]
           n n=0 Mean   SD  Min  0.05   0.1 0.25  0.5 0.75  0.9 0.95  Max
Overall  345    0 1.82 0.905 0.46 0.812 1.010 1.38 1.68 2.09 2.58 3.05 8.90
 - ord   180    0 1.85 0.657 0.69 1.280 1.380 1.55 1.75 2.01 2.38 2.73 8.53
 - dirt   27    0 1.17 0.524 0.46 0.653 0.726 0.91 1.07 1.31 1.56 2.04 3.13
 - scrap  63    0 1.99 1.470 0.67 0.791 0.842 1.21 1.68 2.19 3.17 4.07 8.90
 - wire   75    0 1.86 0.826 0.62 0.784 0.954 1.25 1.82 2.25 2.86 3.58 4.64


Parameter: Intrument height + calculated anomaly depth (m) [ihpd]
           n n=0 Mean   SD  Min 0.05  0.1 0.25  0.5 0.75  0.9 0.95  Max
Overall  345    0 2.49 0.916 0.99 1.28 1.41 1.81 2.38 3.03 3.68 4.10 6.49
 - ord   180    0 2.70 0.812 0.99 1.45 1.70 2.16 2.57 3.24 3.75 4.14 5.47
 - dirt   27    0 2.10 1.010 1.20 1.23 1.31 1.65 1.94 2.26 2.78 3.21 6.46
 - scrap  63    0 2.37 1.040 1.08 1.21 1.30 1.47 2.18 3.01 3.40 4.49 5.94
 - wire   75    0 2.24 0.893 1.17 1.22 1.34 1.55 2.01 2.79 3.38 3.66 6.49


Parameter: Analytic signal approximation (nT/m) [as.approx]
           n n=0 Mean   SD   Min 0.05  0.1 0.25   0.5 0.75   0.9   0.95   Max
Overall  345    0 26.9 54.0 1.26 2.44 3.02 4.26  8.68 27.5  64.0 101.0 561.0
 - ord   180    0 16.8 25.3 1.26 2.42 2.82 3.92  6.63 14.8  46.1  62.2 147.0
 - dirt   27    0 15.4 22.0 1.83 3.28 3.48 4.28  7.09 13.3  37.5  64.5  93.5
 - scrap  63    0 42.8 70.8 1.64 2.45 3.41 7.03 17.40 45.9 103.0 180.0 423.0
 - wire   75    0 41.7 83.2 1.62 2.60 3.24 5.69 16.40 36.7  89.2 133.0 561.0


Parameter: Ratio, as.grid : as.approx [rat.asg.asa]
           n n=0 Mean   SD    Min   0.05   0.1  0.25   0.5 0.75  0.9 0.95  Max
Overall  345    0 1.220 0.903 0.0957 0.125 0.173 0.443 1.100 1.80 2.46 2.78 4.91
 - ord   180    0 1.380 0.944 0.0957 0.137 0.179 0.460 1.370 1.98 2.64 2.88 4.91
 - dirt   27    0 0.859 0.461 0.1050 0.141 0.298 0.591 0.877 1.11 1.25 1.26 2.35
 - scrap  63    0 1.160 1.000 0.1190 0.137 0.169 0.430 0.846 1.76 2.38 3.36 4.19
 - wire   75    0 1.040 0.754 0.1110 0.124 0.134 0.362 1.030 1.45 2.22 2.49 2.87
```

Appendix 3. Publications

Poster Abstract

Partners in Environmental Technology Technical Symposium and Workshop, Nov. 27-29, 2001, Washington D. C., Sponsored by SERDP and ESTCP

Spatial Point Pattern Statistical Models and Optimal Survey Design for Rapid UXO Site Characterization

Dr. William E. Doll
Environmental Sciences Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6038
(865)-576-9930, d8e@ornl.gov

Co-Performers:

Dr. George Ostrouchov, CSMD, ORNL
Mr. T. Jeffrey Gamen, ESD, ORNL
Dr. D.K. Butler, CEERDC
Dr. Max D. Morris, Iowa State Univ.

Modern sensor array technologies are being used to fully characterize large potential UXO sites around the world. The cost per acre for these array-based surveys is often an order of magnitude better than conventional ground surveys. However, the shear size of the problem is such that the volume of work is too great. Statistically valid sampling approaches must be developed to further reduce the survey footprint.

ORNL has recently developed techniques for statistical characterization of UXO contamination that are based on spatial stochastic models for point pattern data. There are two kinds of spatial data. Geostatistical data are quantities that can be measured at any *point* of a given region of interest. Examples of such data are magnetic or electromagnetic ground response signal or concentration of a soil contaminant. Point pattern data are quantities that are obtained by surveying an *area* within a region of interest. Examples of such data are UXO locations or tree locations. The UXO characterization problem begins with geostatistical data (magnetic or EM response), which is then converted to point pattern data (anomaly locations). Here we report on preliminary methods to delineate contaminated areas through contamination intensity estimation by statistical point process models. Results of this method of contamination intensity estimation are presented for data from the Badlands Bombing Range in South Dakota as well as for simulated data.

Point Process Analysis of Geophysical Data for Characterization of UXO Sites

Dr. George Ostrouchov
Oak Ridge National Laboratory
PO Box 2008
Oak Ridge, TN 37831

Co-Performers:
> Dr. W. E. Doll, Dr. D. Wolf, Mr. T. J. Gamey, Dr. Les P. Beard, ORNL
> Dr. M. Morris, Iowa State University
> Dr. D. K. Butler, ERDC

Characterization of sites potentially contaminated with UXO has often used SiteStats/GridStats and UXO Calculator methodology. Although better tools are not readily available, these tools have been shown to have some serious drawbacks including unrealistic assumptions, arbitrary stopping rules, and absence of spatial information. We report on a project that addresses methods for spatial statistical characterization of a site based on samples of geophysical measurements. We emphasize rigorous assessment of uncertainty that is present in the spatial characterization. Our approach is to apply model-based methods for Poisson count data to point patterns derived from geophysical data. Bayesian estimation of all model parameters from the count data provides predictions for areas not sampled along with a complete distribution estimate for each pixel. The estimation and prediction proceeds via Markov chain Monte Carlo and requires substantial computation that can be completed in reasonable time with current PC technology. The computation allows us to make fewer and more realistic assumptions to generate uncertainty estimates.

Our conceptual site model emphasizes three independent sources of correlation: instrument response correlation (electromagnetic and/or magnetic response for a single ordnance scale), ordnance placement correlation (single target, multiple ordnance scale), and target placement correlation (site or multiple target scale). We represent information in our estimates by three components that roughly correspond to the correlation scales: the ordnance list (OL), the ordnance intensity map (OIM), and the target intensity map (TIM). We recommend two or more iterations where an initial site model based on an Archive Search Report and related information is updated with optimally designed samples until a site-specific criterion is met.

Survey design methods are incorporated in this project. The projected performance of geophysical sensors and platforms is considered in development of the first survey, and dig results from each iterative survey are used to refine statistical parameters (e.g. ROC curves) for improved accuracy in the OIMs and TIMs. Information about topography, vegetation, and geology can also be used in the survey design for sample location and selection of sampling technology.

Extended abstract, presented at 2003 Symposium on the Application of Geophysics to Engineering and Environmental Problems

Rapid screening of large-area magnetic data for unexploded ordnance
L.P. Beard, D.A. Wolf, B. Spurgeon*, T.J. Gamey, W.E. Doll
Oak Ridge National Laboratory, Oak Ridge, TN
*Geosoft Europe Ltd., Wallington, Oxfordshire, UK

Abstract

Airborne magnetic surveys can cover hundreds of hectares with very close sensor spacing in a single day. Over unexploded ordnance (UXO) contaminated areas this can translate to thousands of anomalies. Any tool that permits one to rapidly classify anomalies as probable non-UXO and probable UXO is useful. Several geophysical characteristics can be exploited to sort the anomalies, among them signal amplitudes, estimated source depth, and indicators of magnetic remanence. We have developed a grid-based technique that combines information from the total field residual anomaly, the analytic signal, and sensor height to estimate source depth and remanent magnetization. We can then use these and other indicators in statistical schemes to predict whether the source of an anomaly is or is not ordnance.

Introduction

The problem of clearance of UXO from current or former military gunnery or bombing ranges requires a thorough knowledge of where the ordnance is located, and geophysical methods, particularly magnetic and electromagnetic methods, have been widely used for mapping ordnance. Airborne geophysics is increasingly becoming accepted as a way to screen areas of hundreds or thousands of hectares for UXO (Doll, Gamey, and Holladay, 2001). Such surveys typically produce thousands or tens of thousands of anomalies that could be produced by UXO. Early on it was recognized that the vast majority of these anomalies are not caused by hazardous ordnance, but by exploded ordnance fragments or from lost or discarded ferrous articles (e.g., tools, wire, vehicle parts). Methods that reliably discriminate ordnance from non-ordnance can thus save a great deal of time and expense on subsequent cleanup by reducing the number of items to be investigated. To date, model-fitting methods such as the dipole-fitting approach used in the MTADS DAS software (Nelson and McDonald, 1999) have been successfully applied to ordnance discrimination. However, the software in its current state requires the user to choose one anomaly at a time from a grid of total field magnetic data isolate a zone around it, and then performs the inversion. Although results are generally reliable for isolated anomalies, it is ill suited for dealing with dense UXO concentrations, such as occur in the center of a target. Furthermore, the procedure can quickly become tedious for analysis of the considerable number of anomalies resulting from a low level airborne survey. In this paper we describe two alternate approaches based on statistical analysis by which large airborne data sets can be examined quickly for discrimination of UXO.

Statistically based UXO discrimination

We have only recently begun investigating statistically based discrimination methods, after an analysis of dig results based on data collected at the former Badlands Bombing Range (BBR) in South Dakota showed statistical differences between ordnance and non-ordnance. In no instance was the statistical difference so strong that a single parameter could predict whether the source of an anomaly was UXO or not, but the possibility for discrimination increased as more parameters were considered. We used a routine developed to our specifications by Geosoft to rapidly identify and characterize anomalies above a given threshold from an analytic signal map. From these peaks we identified the associated magnetic field anomaly and sensor altitude, and computed a number of parameters that could be used directly or otherwise combined as statistically relevant predictors. From this point we used two different approaches for discrimination—a univariate and a multivariate methods.

Univariate method

What we call the univariate method relies on correlations from dig results based on airborne magnetic data collected at two different sites: an East Coast site and BBR. Both sites were geologically 'clean' in that neither contained basaltic rock or magnetic soils that could complicate any interpretations. We chose six parameters showing correlation with known UXO, and at each anomaly location evaluated whether the parameters fell within the range of the majority of known measured UXO. Each of the six parameters was scored zero if the parameter fell outside a specified range, and one if it fell within the range. For example, almost all ordnance in our known sample pool yielded peak-to-peak magnetic anomalies between 1.0 and 80 nT. Any anomaly falling outside this range was scored zero, as non-UXO. The six characteristics were scored and summed, so that items could have a sum total ranging from 6 (all characteristics in the range of UXO) or zero (all characteristics outside the range for UXO). The six parameters used in the univariate analysis were analytic signal amplitude, magnetic anomaly peak-to-peak magnitude, the distance between the magnetic anomaly peak and low, the ratio of the positive magnetic anomaly lobe to the peak-to-peak magnitude, the estimated source depth, and the angle between magnetic north and the line connecting the positive and negative lobes of the magnetic anomaly (denoted theta).

Multivariate method

Multivariate analysis should provide more information than the univariate approach described above so long as some or all of the variables are correlated, and if the number of known samples is large enough to obtain reliable statistics. The parameters must also be appropriately normalized to remove the effects of different magnitudes for the given parameters. We derived a vector of standard mean parameters μ_0 from a set of measurements over known ordnance items, and compute the symmetric covariance matrix S from the covariances computed for the different variable combinations. The statistical similarity between the known ordnance and the parameter vector x associated with an unknown is given by the Mahalanobis distance (Swan and Sandilands, 1995)

$$D = \{(x - \mu_0)^T \, S^{-1} \, (x - \mu_0)\}^{1/2}.$$

(1)

The smaller the Mahalanobis distance the more closely the unknown resembles ordnance from the known pool of items. The vectors x and μ_0 each have five entries: analytic signal peak, the magnitude of the negative lobe of the magnetic anomaly, the ratio of the positive magnetic anomaly lobe to the peak-to-peak magnitude, the ratio of the distance between the magnetic anomaly positive peak and the analytic signal peak to the instrument height added to the estimated source depth, and theta, as described in the univariate section. The differences in the variables used in the two methods of analysis occurred because the univariate analysis was done prior to a more complete statistical look at the data led to the multivariate approach.

Field results

The methods were applied to a data set acquired by Oak Ridge National Laboratory (ORNL) at an artillery test range in the continental U.S. At one site ordnance was buried for instrument calibration purposes in an area not used as a firing range, and we were given information on ordnance type and location. Descriptions of the ten ordnance items are given in Table 1. Figure 1 shows the residual magnetic field anomaly from a low level survey, flown at a nominal 1.5 m sensor height above ground level using a system developed at ORNL (Doll, Gamey, and Holladay, 2001). Overlain and marked by '+' symbols are locations of ten inert, intact ordnance items commonly found at the test ranges. The smallest items were 60 mm illumination rounds (items 1, 2), followed by 81 mm shells (items 3, 4), 2.75 in rockets (items 5, 6), 105 mm shells (items 7, 8), and the largest targets, 155 mm shells (items 9, 10). The ordnance was buried in two rows, with the items on the west (left) side having an east-west orientation for the long axis of the UXO, and the east (right) row oriented north-south. Shown in Figure 2 is the analytic signal map derived from the magnetic residual. For the 73 analytic signal anomalies in the entire map area at or above 2.0 nT/m, univariate and multivariate statistical analyses were applied. The circle symbols represent the 22 anomalies that were chosen using the univariate classification system as being in the top two categories of probable UXO, i.e. anomalies in which at least five of the six UXO predictors were positive for UXO. Of these 22 anomalies predicted to be UXO, eight occur at or very near the known UXO. Item 2, a 60 mm shell, did not produce enough of an anomaly to register above the 2-nT/m cutoff. The anomaly from item 6, a 2.75 in rocket, is obscured by the large backgrounds anomaly of unknown origin. The triangle symbols represent the same number of items (22) ranked at the top of the multivariate analysis list as most probable UXO. This method ranked nine of the ten known UXO items in the top 22 candidates. The only item missed was item 2, which as previously mentioned, failed the 2.0 nT/m analytic signal cutoff, and so was not included in the pool for statistical analysis. Both methods chose other items as probable UXO as well, and the methods coincided on six of these choices. These anomalies have not been investigated, so we do not know what their sources are. Possibly a few are UXO, but as this site was not in a test firing range, it is more likely they represent non-ordnance: scrap metal, lost tools, or infrastructure.

Table 1. Ordnance description and results of statistical predictions.

Item Number	Ordnance description	Long axis orientation	Univariate prediction	Multivariate prediction
1	60-mm round	E-W	UXO	UXO
2	60-mm round	N-S	undetected	undetected
3	81-mm shell	E-W	UXO	UXO
4	81-mm shell	N-S	UXO	UXO
5	2.75 in rocket	E-W	UXO	UXO
6	2.75 in rocket	N-S	Non-UXO	UXO
7	105-mm shell	E-W	UXO	UXO
8	105-mm shell	N-S	UXO	UXO
9	155-mm shell	E-W	UXO	UXO
10	155-mm shell	N-S	UXO	UXO
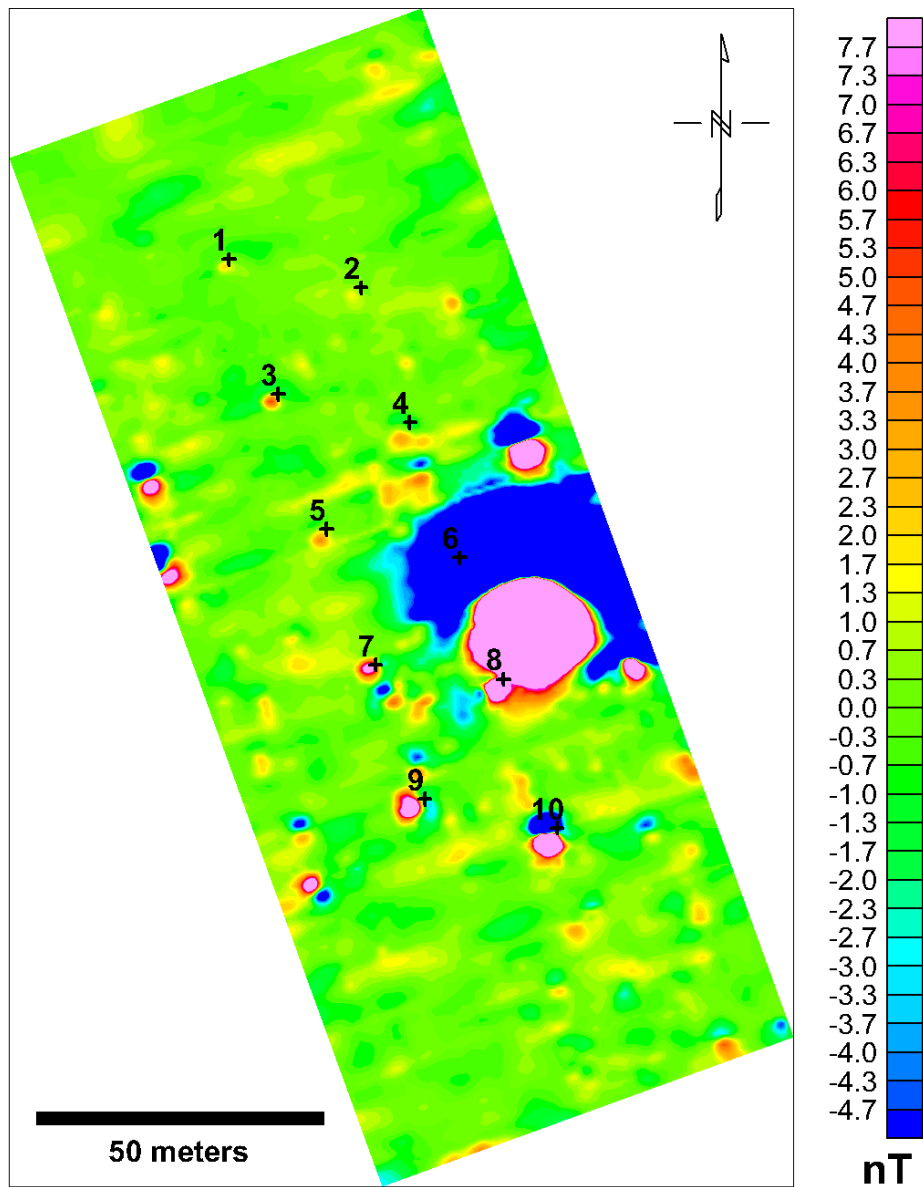
Figure 1. Magnetic anomaly map of helicopter data over calibration grid with known ordnance items overlain.
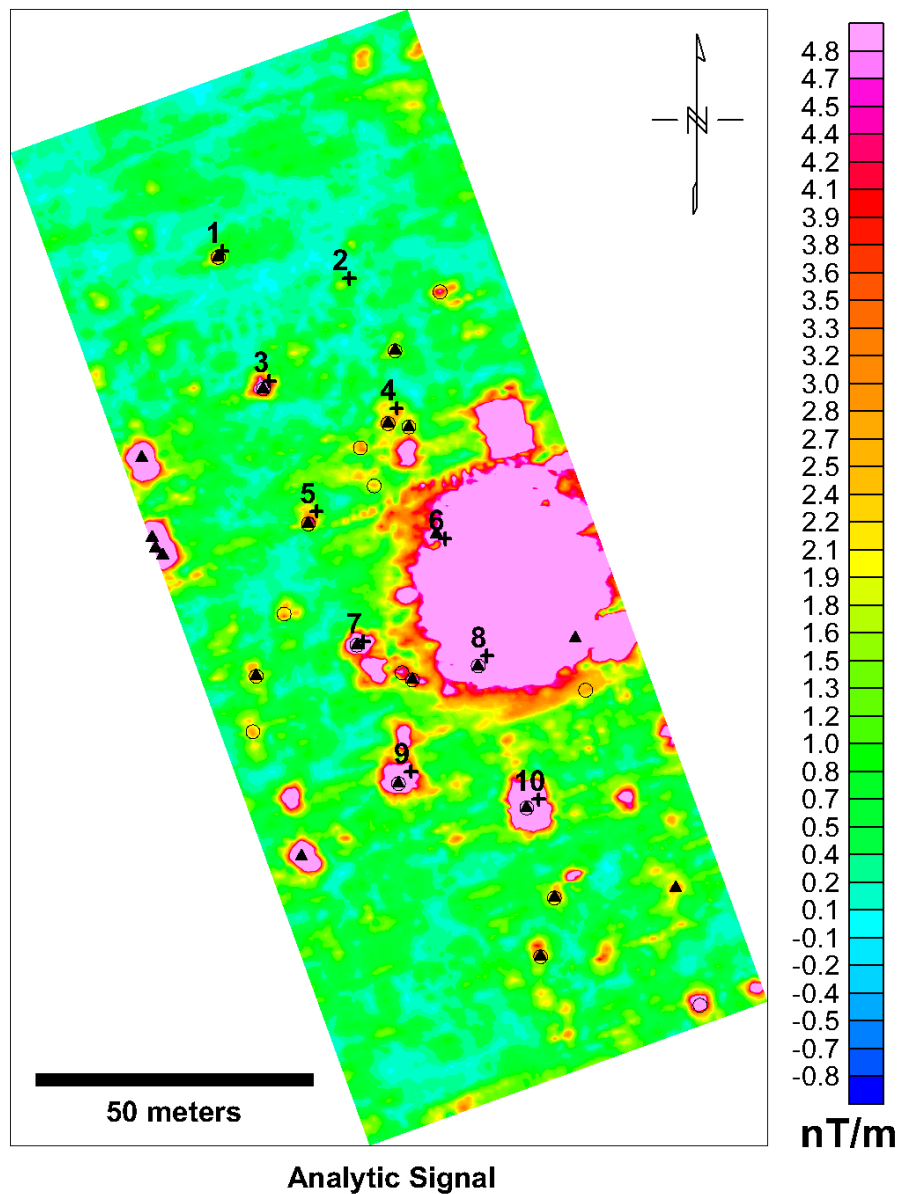
Figure 2. Analytic signal map of calibration grid showing ordnance items and predicted UXO from the univariate approach (circles) and the multivariate approach (triangles).

Discussion and Conclusions

It is difficult to reach any firm conclusion regarding the efficacy of statistical ordnance discrimination at this early stage of work. We have used both inverse model fitting and univariate and multivariate statistical methods for assigning dig locations at several different sites, but at the present time digging has not been done at some sites, and results from digs at the other sites have not been made available as yet. From areas where known items have been buried, each of the three methods reliably predicts known

ordnance to be ordnance, but also predicts some of the unknown anomalies to be ordnance as well. It is distinctly possible that some of these anomalies are produced by ordnance, but we do not yet have information with which to address this matter. The primary difference between the dipole fitting approach in its current state and the statistical approaches is the number of anomalies picked. Our statistical methods pick anomalies from analytic signal peaks that exceed a certain threshold, whereas using the dipole-fitting algorithm the interpreter must choose likely candidates from total field magnetic data. The statistical methods pick on the order of ten times more anomalies for evaluation than is typically chosen by an interpreter using model fitting because so many of the anomalies appear too weak or are insufficiently isolated for inversion. If the highly ranked items on the dig lists produced by statistical methods prove to be mostly real ordnance, then more is good. As dig results become available, we should be able to improve our pool of characteristics of known ordnance items and improve the reliability of the statistical methods.

Acknowledgments

References

1. Doll, W.E., Gamey, T.J., and Holladay, J.S., 2001, Current research into airborne UXO detection: SAGEEP 2001 Proceedings, Denver, Colorado, on CD-ROM.

2. Nelson, H.H. and McDonald, J.R., 1999, Target shape classification using the MTADS: UXO/Countermine Forum 1999 Proceedings, Alexandria, Virginia, on CD-ROM.

3. Swan, A.R.H. and Sandilands, M., 1995, Introduction to Geological Data Analysis, Blackwell Science, 446 pp.